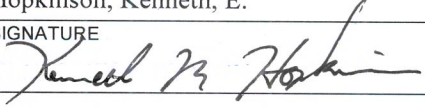
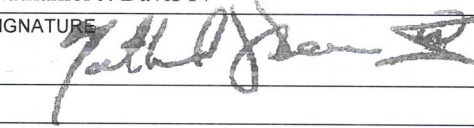


88th ABW PUBLIC AFFAIRS SECURITY AND POLICY REVIEW WORKSHEET

NOTE: Public release clearance is NOT required for material presented in a closed meeting and which will not be made available to the general public, on the Internet, in print or electronic media.

AFITGCOENG1218

1. SUBMITTER NAME janice.jones@afit.edu		OFFICE SYMBOL ENG	PHONE 56565	2. DATE SUBMITTED
3. AUTHOR(S) NAME Kasperek, Andrew, T.		ORGANIZATION AFIT		PHONE 56565
4. DOCUMENT TITLE Enhancing Trust in Smart Grid by Applying a Mod EWMA Alg		6. CONFERENCE/EVENT/PUBLICATION NAME N/A		7. DATE NEEDED N/A
5. DOCUMENT TYPE <input type="checkbox"/> ABSTRACT <input type="checkbox"/> TECH REPORT <input type="checkbox"/> JOURNAL ARTICLE <input type="checkbox"/> VIDEO <input type="checkbox"/> SPEECH <input type="checkbox"/> TECH PAPER <input type="checkbox"/> BRIEFING CHARTS <input type="checkbox"/> PHOTO <input checked="" type="checkbox"/> THESIS/DISSERTATION <input type="checkbox"/> OTHER _____		8. EVENT LOCATION N/A		9. EVENT DATE N/A
10. RELATED CASES PREVIOUSLY APPROVED		11. OTHER AGENCY COORDINATION REQUIRED (List contact information)		12. OTHER INFORMATION
13. NATIONAL SECURITY STATUTES/TECHNOLOGY ISSUES a. <input type="checkbox"/> YES <input checked="" type="checkbox"/> NO Are any aspects of this technology included in: U.S. Munitions List; ITAR 22, CFR Part 121; CCL; Security Classification Guide; DD Form 254 or a Technology Protection Plan? (If YES, please explain in Block 16) b. <input checked="" type="checkbox"/> YES <input type="checkbox"/> NO Does this information meet the criteria for Distribution Statement "A" - unclassified, unlimited distribution?		14. NATIONAL SECURITY STATUTES/TECHNOLOGY ISSUES a. <input type="checkbox"/> YES <input checked="" type="checkbox"/> NO If this material results from an international agreement is the USAF authorized to release program information? (If NO, please identify release authority organization) b. <input type="checkbox"/> YES <input checked="" type="checkbox"/> NA If this is a joint program, does your organization maintain primary management responsibility and authority to release all information? (If NO, please provide name of lead organization/POC (i.e., DARPA, NAS, ARMY, etc.) c. <input type="checkbox"/> YES <input checked="" type="checkbox"/> NA If this information is for a SBIR contractor and the program manager is responsible for public release, is a waiver letter on the file granting permission to release?		
15. BUDGET CATEGORIES (Funding is under Budget Category-Program Element) <input type="checkbox"/> 6.1. <input type="checkbox"/> 6.3. <input checked="" type="checkbox"/> N/A <input type="checkbox"/> 6.2. <input type="checkbox"/> 6.4./HIGHER <input type="checkbox"/> OTHER _____				
16. EXPLANATION				
17. ORIGINATOR I certify the attached material is unclassified, technically accurate, contains no critical military technology, is not subject to export controls and is suitable for public release.		PRINT NAME Kasperek, Andrew, T. SIGNATURE KASPEREK.ANDREW.T.1058536187, DATE (YYYYMMDD) 20120523		
18. TECHNICAL REVIEW AND CERTIFICATION I certify the information contained in the attached document is technically accurate; does not disclose classified, sensitive, or military critical technology, does not violate proprietary rights, copyright restrictions; and is not subject to export control regulations. I further certify that this information is suitable for public release.				
		PRINT NAME Hopkinson, Kenneth, E. SIGNATURE  DATE (YYYYMMDD) 20120523		
19. SECURITY MANAGER REVIEW Signature certifies that the information has been reviewed and the information contains no Operational Security issues.				
		PRINT NAME Nathaniel J. Davis IV SIGNATURE  DATE (YYYYMMDD) 20120615		
20. ADDITIONAL REVIEW I certify that this information is suitable for public release.				
		PRINT NAME DATE (YYYYMMDD) SIGNATURE CLICK HERE TO SIGN		
21. PA USE ONLY				
<input type="checkbox"/> APPROVED <input type="checkbox"/> AS AMENDED <input type="checkbox"/> DISAPPROVED	DATE	PAO SIGNATURE		CASE NUMBER



ENHANCING TRUST IN THE SMART GRID BY APPLYING A
MODIFIED EXPONENTIALLY WEIGHTED MOVING AVERAGES
ALGORITHM

THESIS

Andrew T. Kasperek, Captain, USAF

AFIT/GCO/ENG/12-18

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, the Department of Defense, or the United States Government.

This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States

AFIT/GCO/ENG/12-18

ENHANCING TRUST IN THE SMART GRID BY APPLYING A
MODIFIED EXPONENTIALLY WEIGHTED MOVING AVERAGES
ALGORITHM

THESIS

Presented to the Faculty
Department of Electrical and Computer Engineering
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
in Partial Fulfillment of the Requirements for the
Degree of Master of Science

Andrew T. Kasperek, B.S.C.S., M.I.S.
Captain, USAF

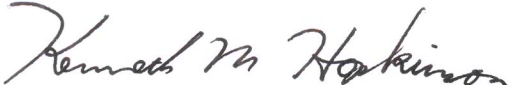
June 2012

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

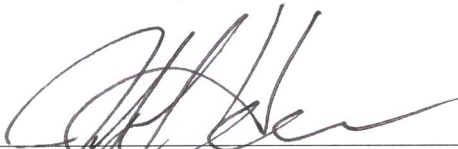
ENHANCING TRUST IN THE SMART GRID BY APPLYING A
MODIFIED EXPONENTIALLY WEIGHTED MOVING AVERAGES
ALGORITHM

Andrew T. Kasperek, B.S.C.S., M.I.S.
Captain, USAF

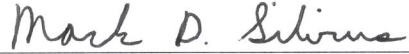
Approved:


Kenneth M. Hopkins, PhD (Chairman)

23 May 2012
Date


Jeffrey M. Hemmes, Lt Col, PhD (Committee Member)

29 May 2012
Date


Mark D. Silvius, Maj, PhD (Committee Member)

29 May 2012
Date

Abstract

The main contribution of this thesis is the development and application of a modified Exponentially Weighted Moving Algorithm (EWMA) algorithm, and its ability to robustly function in the face varying numbers of *bad* (malicious or malfunctioning) Special Protection System (SPS) nodes. Simulation results support the use of the proposed modified EWMA reputation based trust module in SPSs within a smart grid environment. This modification results in the ability to easily maintain the system above the minimum acceptable frequency of 58.8 Hz at the 95% confidence interval, when challenged with test cases containing 5, 10 and 15 *bad* node test cases out of 31 total load nodes.

These promising results are realized by incorporating the optimal modified EWMA strategy, as identified by Receiver Operating Characteristic (ROC) techniques, where an optimal strategy is revealed. The optimal strategy maximizes true positives while minimizing false positives.

Implementation of a modified EWMA within a reputation based special protection system does not account for each scenario that an electrical power engineer may face in the field. Instead, this research demonstrates that such an algorithm provides a robust environment to test within, in the hope of successfully meeting challenges and/or opportunities of the future.

For my wife and children, for their numerous sacrifices throughout my career.

Acknowledgments

Foremost, I would like to express deep appreciation and gratitude to my faculty advisor, Dr. Kenneth Hopkinson, for his insights and levity during this thesis effort. I would also like to thank my committee members, Lt Col Jeffrey Hemmes and Maj Mark Silvius, for their patience and advice. Additionally, special thanks to Dr. Rusty Baldwin, Dr. Timothy Lacey, Ms. Kathy Atkinson and Ms. Janice Jones for their steadfast motivation and continually open doors. Finally, special recognition is deserved for my comrades in the Cyber Advanced Networking in Mobile Applications Laboratory (ANiMaL), especially Maj José Fadul, Capt Jason Bindewald, Capt Joan Betances, Capt Tony Simpson and MSgt Crystal Shipman... You entered these halls as fellow students, and we now part as friends.

Andrew T. Kasperek

Table of Contents

	Page
Abstract	iv
Dedication	v
Acknowledgments	vi
List of Figures	x
List of Tables	xii
List of Symbols	xiii
List of Abbreviations	xiv
 1 Introduction	 1
1.1 Overview	1
1.2 Background	2
1.3 Problem Statement	3
1.4 Research Goals	4
1.5 Contributions	4
1.6 Chapter Review	4
 2 Literature Review	 6
2.1 Introduction	6
2.2 Critical Infrastructure	6
2.3 Governance	8
2.4 SCADA	10
2.4.1 Overview	10
2.4.2 Architecture	11
2.4.3 Operation	11
2.4.4 Threats and Vulnerabilities	12
2.5 Electrical Power Generation	13
2.5.1 Characteristics	13
2.5.2 Operation	14
2.6 Smart Grid	16
2.6.1 Attributes	17
2.6.2 Threats and Vulnerabilities	17
2.7 Special Protection System	18

2.7.1	Traits	20
2.7.2	Power System Stability	20
2.7.3	Generation Rejection Scheme	22
2.7.4	Underfrequency Load Shedding	22
2.7.5	Trust Management	23
2.7.6	Reputation Based Trust	23
2.7.7	Special Protection System Trust Module	24
2.7.8	Greedy Sorting Algorithm	25
2.8	Exponentially Weighted Moving Averages	26
2.8.1	Overview	27
2.8.2	Sample Calculations	27
2.8.3	Traditional EWMA Example	28
2.8.4	SPS Applicability	30
3	Methodology	32
3.1	Overview	32
3.2	Modified Exponentially Weighted Moving Average	32
3.2.1	Analysis	36
3.3	Testing Environment	38
3.3.1	EPOCHS Simulation Environment	38
3.3.2	Network Simulator 2	39
3.3.3	PSS/E	40
3.3.4	AgentHQ	40
3.3.5	RTI	41
3.3.6	Component Interaction	41
3.4	System Studied	41
3.4.1	Special Protection Scheme Action Goal	42
3.5	Test Scenario	43
3.6	Experimental Design or Test Cases	45
3.6.1	Stage 1 (Normality Testing)	45
3.6.2	Stage 2 (Modified EWMA)	46
3.6.3	Stage 3 (5, 10 and 15 Bad Node Frequencies)	46
3.7	Analysis	47
3.8	Methodology Summary	47
4	Analysis and Results	48
4.1	Overview	48
4.2	Stage 1 (Normality Testing)	48
4.3	Stage 2 (Modified EWMA)	51
4.4	Stage 3 (5, 10 and 15 Bad Node Frequencies)	53
4.5	Analysis and Results Summary	57

5	Conclusion and Future Work	58
5.1	Chapter Overview	58
5.2	Conclusions of Research	58
5.3	Recommendations for Future Research	59
	Appendix: Receiver Operating Characteristic Data	62
	Bibliography	64
	Vita	69

List of Figures

Figure	Page
2.1 Power Production and Distribution System	14
2.2 Steam Turbine Partial or Full Load Operating Limitations During Abnormal Frequency	15
2.3 Traditional EWMA Graph on Data Set Y , with $\lambda = 0.6$	30
3.1 Modified EWMA Graph on Data Set Y , with $\lambda = 0.6$	35
3.2 Confusion Matrix	36
3.3 Basic ROC Graph Showing Five Discrete Classifiers	37
3.4 Relationship Between EPOCHS Components	38
3.5 NS2 Architecture	39
3.6 Representation of a Smart Grid Wide Area Network	40
4.1 Histogram Generated from Simulation Results for Traditional SPS with 5 Bad Nodes)	49
4.2 Quantile-Quantile Plot Generated from Simulation Data for Traditional SPS with 5 Bad Nodes	50
4.3 Receiver Operating Characteristic Curve Data with Optimal Strategy Identi- fied; Point (0.00, 1.00) with $\lambda = 0.1$ and Trust Threshold = 0.5	51
4.4 Receiver Operating Characteristic Data With Optimal Strategy for λ and Trust Threshold Value Identified	52
4.5 Frequency Observations for Optimally Modified EWMA with 15 Bad Nodes at 95% Confidence Interval	54
4.6 Frequency Observations for Traditional EWMA Trust Implementation ($\lambda = 1$, Trust Threshold = 0.5) with 15 Bad Nodes at 95% Confidence Interval	54
4.7 Frequency Observations for No Trust with 15 Bad Nodes at 95% Confidence Interval	55

4.8	Comparison of Final Frequency Values for 5, 10 and 15 Bad Nodes at 95% Confidence Interval	56
5.1	Comparison of 5, 10, 15-21 Bad Node Test Cases with Optimal Trust Module 95% Confidence Interval	60

List of Tables

Table	Page
2.1 Eight Critical Infrastructures According to E.O. 13010	7
2.2 GAO-Identified Challenges to Securing Smart Grid Systems	10
2.3 Sources and Motivations for Utility Disruptions and Attack	12
2.4 Percentages of Most Common SPS Types	19
2.5 Sorted Nodes For Possible Load Shedding	25
2.6 Intermediate Calculations of a Traditional EWMA Algorithm	29
3.1 Intermediate Calculations of Modified EWMA Algorithm	35
A.1 Receiver Operating Characteristic Data Used to Determine Appropriate λ and Trust Threshold Values (1 of 2)	62
A.2 Receiver Operating Characteristic Data Used to Determine Appropriate λ and Trust Threshold Values (2 of 2)	63

List of Symbols

Symbol	Page
Hz hertz	14
MW megawatt	25
λ Exponentially Weighted Moving Average (EWMA) Smoothing Factor	26
Z_i Original EWMA Algorithm	28
Z_0 Average of Monitored Process using Traditional EWMA	28
\hat{Z}_{N+1} Average of Monitored Process using Modified EWMA	34
\hat{Z}_j Modified EWMA Algorithm	34
kV kilovolt	41

List of Abbreviations

Abbreviation	Page
ICS	Industrial Control Systems 1
SCADA	Supervisory Control and Data Acquisition 1
CI	Critical Infrastructure 3
EISA	Energy Independence and Security Act of 2007 3
FERC	Federal Energy Regulatory Commission 3
PDD	Presidential Decision Directive 3
SPS	Special Protection System (also known as Special Protection Scheme) 3
i.e.	id est 3
EWMA	Exponentially Weighted Moving Averages 6
EO	Executive Order 6
NSTAC	National Security Telecommunications Advisory Committee 8
IT	Information Technology 9
NIST	National Institute of Standards and Technology 9
PLC	Programmable Logic Controllers 10
HMI	Human Machine Interface 11
RTU	Remote Terminal Unit 11
e.g.	exempli gratia 11
HVDC	High Voltage Direct Current 19
etc.	et cetera 21
GRS	Generator Rejection Scheme 22
TP	True Positive 36
FN	False Negative 36
TN	True Negative 36
FP	False Positive 36

TPR	True Positive Rate	37
FPR	False Positive Rate	37
COTS	Commercial Off-The-Shelf	38
IED	Intelligent Electronic Device	39
Q-Q	Quantile-Quantile	46
ROC	Receiver Operating Characteristics	47
ANOVA	Analysis of Variance	47

ENHANCING TRUST IN THE SMART GRID BY APPLYING A MODIFIED EXPONENTIALLY WEIGHTED MOVING AVERAGES ALGORITHM

1 Introduction

This chapter provides a general introduction of the thesis subject area in a general way and an overview of the problem. The importance and motivation of the problem addressed is also presented. Finally, this chapter outlines research goals and the structure of research contained in this theses and also an overview of the remaining parts of the thesis itself.

1.1 Overview

Identified as one of our nation's critical resources [14], the electric power grid is vital not only to the national security of the Unites States, but also its way of life. President Obama astutely summarized a precautionary tale to those that take on today's technological challenges:

It's the great irony of our Information Age—the very technologies that empower us to create and to build also empower those who would disrupt and destroy [45].

The electric power system falls under the broad umbrella of Industrial Control Systems (ICS), and are managed through the use of Supervisory Control and Data Acquisition (SCADA) equipment and processes, with and it associated attributes, both positive and negative. It is important not to frame any discussion on improving

performance and reliability of the electrical power system with these considerations in mind, but to also ensure that security and robustness are key attributes, due to the critical nature of this man-made resource.

1.2 Background

The commercial use of electricity began in the late 1870s when arc lamps were used for lighthouse illumination and street lighting, and the first complete electric power system (comprising a generator, cable, fuse, meter and loads) was built by Thomas Edison in September 1882 [39]. Today, almost all of the utilities in the United States and Canada are part of one very large and enormously complex interconnected system.

Complicating the issue is the fact that the electric power grid is not owned nor managed by one single organization. Rather, it is the conglomeration of numerous companies, usually regional in nature, interconnected to allow power to flow to your home. Key to understanding the critical infrastructure issues associated with the electric power grid is the ability to frame the grid as one large information system. It is essential that any modifications made to current systems account for emerging technologies, to allow seamless integration now and in the future.

With smart grid technologies taunting new abilities such as energy management and real-time pricing, there has been much recent discussion on what distribution systems of the future can and should look like. That is why it is important to understand the characteristics of the smart grid and how to frame accurately new challenges may be associated with its implementation.

Integrated throughout this security challenge and modernization effort is governance primarily at the federal level. The critical infrastructure discussion began in earnest in the late 1990's, with initial reports to the President addressing security of the electric power control networks and the electrical power grid [49]. The security of the nation's

Critical Infrastructure (CI) was thrust into the forefront following the tragic events of September 11th, 2001.

By 2007, the Energy Independence and Security Act of 2007 (EISA) had tasked the Federal Energy Regulatory Commission (FERC) with specific responsibilities with regard to the adoption of smart grid guidelines and standards [52]. Specific challenges that must be addressed with any smart grid enhancements are identified in Section 2.3.

It is with the basic background identified in this section that this and related research has built simulation environments to help ensure new technology development not only addresses these concerns, but also meet evolving federal regulation standards. As identified in Presidential Decision Directive (PDD)-63, President Clinton's intent at the time was that "The United States will take all necessary measures to swiftly eliminate any significant vulnerability to both physical and cyber attacks on our critical infrastructures, including especially our cyber systems" [14]. The research conducted within this thesis, enhancing the reliability of the electrical power grid's special protection systems, is certainly in alignment with the former President's vision and should be in the mind of any similar researcher. This high-level background helps frame the problem statement that this research hopes to address.

1.3 Problem Statement

An incorrect decision made by a electric power grid Special Protection System (SPS), also known as Special Protection Scheme, can have drastic consequences and result in needless power service interruption. Current protection system methodology with regard to load shedding do not accurately determine optimal loads for shedding. Therefore, it is often the case that *good* (i.e., cooperating, non-malicious and/or non-malfunctioning) nodes are incorrectly identified as *untrusted* and *bad* (i.e., not cooperating, malicious and/or malfunctioning) nodes are incorrectly identified as *trusted*.

1.4 Research Goals

The goal of this research is to develop a robust algorithm that will efficiently and accurately calculate reputation based trust within electrical power grid special protection systems, allowing the system to maintain an acceptable frequency level and accurately classify *good* and *bad* nodes. The algorithm will be tunable to specific protection system applications, adjusting for individual characteristics of each application, such as associated background noise.

1.5 Contributions

Research contained in this thesis is driven by a novel approach to determining trust within the electric power grid. Specifically, the creation and subsequent testing of a unique algorithm to determine trust within an SPS promises to not only provide accurate trust calculations, but minimize associated error.

1.6 Chapter Review

This chapter presented the research topic at a very high level, and the remainder of this document is to fully document the research process of this thesis.

Chapter 2 presents an overview of literature that supports the research design and key simulation parameters. This chapter also presents current research in the area.

Chapter 3 comprises the experimental methodology, to include goals and hypothesis, testing environment and associated test cases and an overview of how results gathered will be analyzed. It is within this chapter that the proposed algorithm is presented.

Chapter 4 reports the results from applying the design identified in Chapter 3. Both observational and interpretive analysis techniques are used to convey a clear understanding of the impact of this research.

Chapter 5 identifies conclusions and recommendations for future research in this topic area. The conclusions assert whether or not results from this research warrant additional consideration and or immediate implementation. Throughout the course of this research, numerous ideas presented themselves as promising research leads, but were simply outside the scope of this thesis. Chapter 5 presents these ideas and frames their significance in the context of expanding the body of work in this topic area.

2 Literature Review

2.1 Introduction

This chapter provides brief introductory material, as well as a review of literature, concepts and current research relevant to the protection of critical infrastructures, the nation's electric power system and the components of associated Supervisory Control and Data Acquisition (SCADA) management systems. Additionally, challenges associated with emerging smart grid technologies and their implementation are presented. Then, an overview of Special Protection Systems (SPS)s, also known as Special Protection Schemes, and current research related to that topic area is addressed. This chapter concludes with a look at traditional Exponentially Weighted Moving Averages (EWMA) implementations and its direct applicability to an SPS.

Integral to this research is governance pertaining to these topic areas, as regulation and oversight guide not only the future of the electric power grid, but also mandate constraints in which evolutions to this topic area must adhere. Therefore, the literature review contained herein begins with an overview of how the electrical power grid fits the definition of Critical Infrastructure (CI) and a survey of relevant governance.

2.2 Critical Infrastructure

All characteristics of the US electric power system, including vulnerabilities, have garnered much deserved attention over recent years, as the system itself clearly falls into the category of critical infrastructure. On July 15, 1996, President Clinton signed Executive Order 13010 establishing the President's Commission on Critical Infrastructure Protection [13]. This Executive Order (EO) defined "infrastructure" as:

The framework of interdependent networks and systems comprising identifiable industries, institutions (including people and procedures), and

distribution capabilities that provide a reliable flow of products and services essential to the defense of government at all levels, and society as a whole.

Table 2.1: Eight Critical Infrastructures According to E.O. 13010 [13]

No.	Government Sector(s)
1.	Telecommunications
2.	Electrical Power Systems
3.	Gas and Oil Storage and Transportation
4.	Banking and Finance
5.	Transportation
6.	Water Supply Systems
7.	Emergency Services
8.	Continuity of Government

E.O. 13010 broadened the list of critical infrastructure sectors to include electrical power system by name as identified in Table 2.1. In 1998, President Clinton continued the evolution of the term critical infrastructure to include those physical and cyber-based systems essential to the minimum operations of the economy and government. They include, but are not limited to, telecommunications, energy, banking and finance, transportation, water systems and emergency services, both governmental and private[14]. Additionally, President Clinton expressly acknowledged that "Many of the nation's critical infrastructures have historically been physically and logically separate systems that had little interdependence".

In a 1996 study, it was determined that over 90 percent of the nation's critical infrastructures were privately owned and operated [34]. This privatization certainly extends to both SCADA systems, and specifically the US electrical power systems.

Therefore, it should be no surprise that there is a wealth of governance guiding the power grid's security and modernization. Of specific interest to this research is governance pertaining to these topic areas, as regulation and oversight guide not only the future of the electric power grid, but also mandate constraints in which evolutions to this topic area must adhere. It is within this context that governance is reviewed.

2.3 Governance

In March 1997, the National Security Telecommunications Advisory Committee (NSTAC) issued a report to the President that assessed the security of the electric power control networks and electric power grid. The report warned of utilities rapidly expanding their use of information systems and interconnecting previously isolated networks because of competition, aging proprietary systems, and reductions in staff and operating margins [49].

Although Commission identified electronic intrusion of the utilities' information systems and networks as an emerging threat, it found that the industry considered the primary threat to information systems to be from insiders. Even though the industry at the time focused much of its attention toward the ever-present insider threat, the NSTAC made the determination that substations presented the most significant information security vulnerability in the power grid due in part to the vulnerabilities associated with widespread use of dial-up modems and the use of public networks.

In early 2001, the NSTAC Information Sharing for Critical Infrastructure Task Force consolidated detailed research and analysis done over recent years, and also requested industry advice and recommendations for revision of the National Plan, in order to provide sound recommendations to the President [50]. They framed the problem very well when they stated, "While Government is focusing on protecting national security, preventing future attacks, and identifying and punishing attackers, private owners of infrastructures are more concerned with common business imperatives. As a result of this dichotomy, any

solution to, or recommendation for, the protection of critical infrastructures require the participation of private industry in concert with the Government.

By 2007, the electric industry's increased incorporation of Information Technology (IT) systems as part of the smart grid effort now garnered congressional concern. evaluating the growing concern that smart grid efforts, if not implemented securely, could cause the electric grid to become more vulnerable to attacks and loss of services.

As a result, the Energy Independence and Security Act of 2007 (EISA) provided the National Institute of Standards and Technology (NIST) and Federal Energy Regulatory Commission (FERC) with specific responsibilities with regard to coordinating the development and adoption of smart grid guidelines and standards.

As the audit, evaluation and investigative arm of the United States Congress, the Government Accounting Office was asked to[52]:

1. Assess the extent to which NIST has developed smart grid cybersecurity guidelines
2. Evaluate FERC's approach for adopting and monitoring smart grid cybersecurity and other standards
3. Identify challenges associated with smart grid cybersecurity

With respect to smart grid systems, GAO identified the six key challenges identified in Table 2.2. To address these challenges, the National Institute of Standards and Technology developed and issued a first version of its smart grid cyberspace guidelines in August of 2010. The agency developed the guidelines for entities such as electric companies involved in implementing smart grids[52]. It is important to note that as the transition to smart grid technologies advances, smart grid data availability places considerably more stringent demands on the communication and control system than traditional SCADA systems do [36]. Therefore, it is necessary to understand the underlying legacy SCADA systems, their current usage in the operation of the electric

Table 2.2: GAO-Identified Challenges to Securing Smart Grid Systems [52]

No.	Specific Challenge
1.	Aspects of the regulatory environment may make it difficult to ensure smart grid systems cybersecurity
2.	The electric industry does not have an effective mechanism for sharing information on cybersecurity
3.	Utilities are focusing on regulatory compliance instead of comprehensive security
4.	Consumers are not adequately informed about the benefits, costs, and risks associated with smart grid systems
5.	There is a lack of security features being built into certain smart grid systems
6.	The electricity industry does not have metrics for evaluating cybersecurity

power system and consideration that must be made when attempting to make modifications.

2.4 SCADA

Industrial Control Systems, such as electric power generation plants, are large, distributed complexes, requiring plant operators to continuously monitor and control many different sections of the plant to ensure its proper operation [34]. This monitoring is accomplished through the use of SCADA systems.

2.4.1 Overview. SCADA is short for Supervisory Control And Data Acquisition, and as the implies, the focus of SCADA is on the supervisory level of operation. It is generally used to control dispersed assets using centralized data acquisition and supervisory controls[34]. As such, it is a purely software package that is positioned on top of hardware to which it is interfaced, in general via Programmable Logic Controllers (PLC), or other commercial hardware modules [18]. Initially, ICSs had little resemblance to traditional IT systems in that ICSs were isolated systems running proprietary control protocols using specialized hardware and software [34]. SCADA is pervasive in the

generation and distribution of energy with each utility and cooperative having its own SCADA system [23].

2.4.2 Architecture. A SCADA system is identified by two basic layers: the "client layer" which enables the man machine interface and the "data server layer" which is responsible for the majority of process data control activities [18].

The SCADA master station consists of the SCADA master servers and the Human Machine Interface (HMI). The master station is located in a central control center from where operators can monitor the entire system. SCADA master servers run the server-side applications that communicate with the Remote Terminal Unit (RTU). The SCADA master servers poll the RTUs for data and send control messages to supervise and control the utility's physical infrastructure. Backup servers are used to increase fault-tolerance of the system [24].

Data servers communicate with devices in the field through PLCs. PLCs are connected to the data servers either directly or via networks or field buses that are proprietary (exempli gratia (e.g.) Siemens H1), or non-proprietary (e.g. Profibus) [18]. The data servers are responsible for data acquisition and handling (e.g. polling controllers, alarm checking, calculations, logging and archiving) on a set of parameters. This pulling of data from remote locations permits operators to monitor and control remote assets and processes in real time.

2.4.3 Operation. To provide real-time data updates from the field, a SCADA system needs remote sensory and communications capabilities. Electronic devices called RTUs are located at each point where measurements are to be taken or where process equipment is to be controlled. The central computer continuously polls the field-based RTU to fetch their current measurement message containing updated values, and repeating

that operation with subsequent RTUs, until all have been processed. That sequence is then repeated over and over, without end [58].

2.4.4 Threats and Vulnerabilities. There are numerous sources and motivations for disruption within a SCADA system as identified in Table 2.3

Table 2.3: Sources and Motivations for Utility Disruptions and Attack [28]

Source	Reason
Industrial sabotage or theft	Financial advantage in insider trading or competing vendor partnerships
Concentrated physical and cyber attack	Destruction, terror or activism
Vendor compromise	Easier to target the supplier than the defended infrastructure itself [26]
Technical design error or environmental influence	Hardware or code; network design, installation and configuration; or interferences from other technologies in the environment
Natural disasters	Earthquakes, tornadoes, volcanoes, fires, thunderstorms and snow storms
Operator error	Misjudgement, misconfiguration, or failure to remember operational details, resulting in dangerous or costly results

Additionally, there have been many real-world incidents affecting SCADA systems, and many others never publicized, that clearly illustrate vulnerabilities [26] [16]:

- During the Cold War, the U.S. provided Trojan firmware to the Soviet Union, causing a pipeline to explode in one of the world's largest non-nuclear explosions [26].
- In 2000, a disgruntled employee rigged a computerized control system at a water treatment plant in Australia, releasing more than 200,000 gallons of sewage into parks, rivers and the grounds of a Hyatt hotel [25]

- In 2001, hackers hacked CAL-ISO, Californias primary power grid operator, and were not discovered for 17 days [11]
- In 2003, the Slammer Worm took Ohios Davis-Besse nuclear plant safety monitor offline for five hours [26]
- In 2008, a senior Central Intelligence Agency official, Tom Donahue, told a meeting of utility company representatives in New Orleans that a cyberattack had taken out power equipment in multiple regions outside the U.S. Mr. Donahue stated that the outage was followed with extortion demands [25]

Cyber attacks on U.S. SCADA networks have the potential to affect supplies of gasoline, electricity or water, ultimately impacting stock prices on a global level [16]. It is now clear that SCADA system in general, and the electric power system specifically, has numerous vulnerabilities and protecting it warrants additional consideration. Therefore, reviewed of power production and distribution system on the U.S. power grid is warranted.

2.5 Electrical Power Generation

Continuous control of electric power generation to match changes in load has been a standing problem which has attracted the attention of the workers and researchers of power operation and control [8]. This is due to the fact that unpredictable changes in load frequently cause power generation-consumption mismatches, adversely affecting the quality of generated power due to the offsetting of the desired frequency value.

2.5.1 Characteristics. Although distribution circuits come in many different configurations and circuit lengths, most share many common characteristics as identified in Figure 2.1. These components include the generation station, generating step up

transformers, the transmission lines, substation step down transformers and then the end customers.

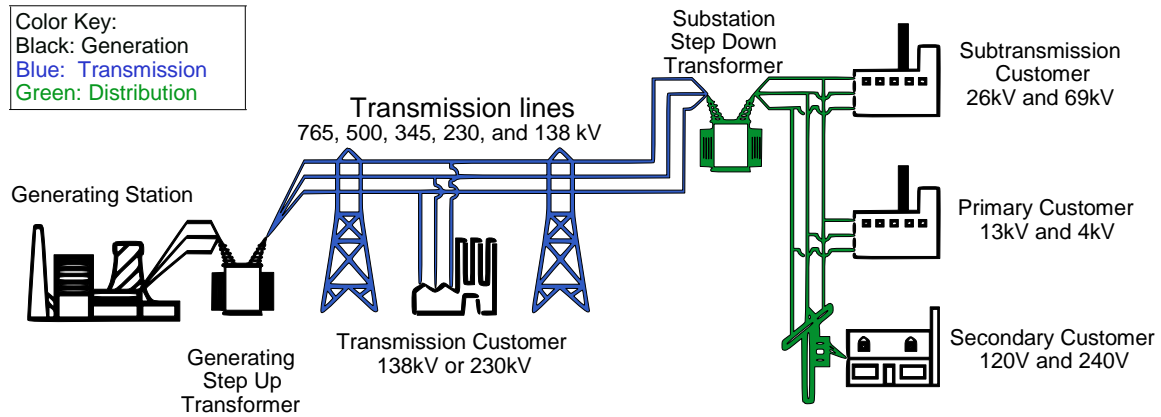


Figure 2.1: Power Production and Distribution System [42]

2.5.2 Operation. Deviation of frequency from its nominal value should be minimized and kept within rigid limits in order for electric power consuming and frequency dependent control equipment to operate satisfactorily [8]. Additionally, turbines used for power production are designed to operate at specific frequencies and incur stress related damage when operating at higher or lower frequencies. Manufacturers often provide abnormal operating characteristics, recognizing that each generation device will have its own unique limits. Figure 2.2 illustrates the operational limits of a representative steam turbine with the following characteristics as measured in Hertz (Hz) [8]:

- The areas between 59.5 Hz and 60.5 Hz are areas of unrestricted time operating frequency limits
- Operation between 58.5 Hz and 57.9 Hz is permitted for ten minutes before turbine blade damage is probable

If a unit operates within this frequency band for one minute, then nine more minutes of operation within this band are permitted over the life of the blade [7]

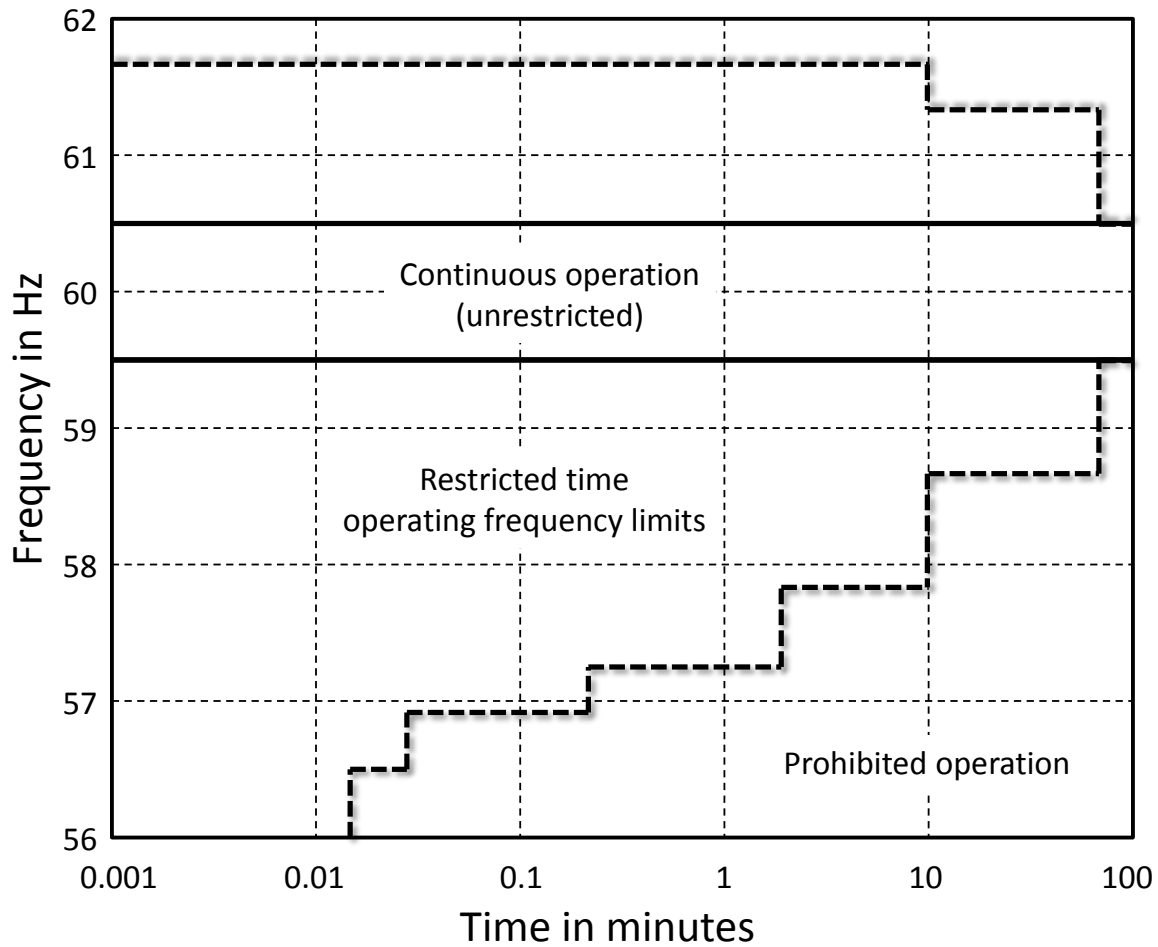


Figure 2.2: Steam Turbine Partial or Full Load Operating Limitations During Abnormal Frequency [1]

With regard to operational limits, it is important to remember that time spent in a given frequency band is cumulative and, for preventative maintenance purposes, is usually considered independent of the time accumulated in any other band. Since fatigue life is used up during abnormal underfrequency operation, the time spent in an underfrequency event should be minimized as much as possible. It is with these operational limit

considerations in mind that several researchers have created a *minimum acceptable frequency* of 58.8 Hz [32] [21] [54], as operating below this frequency threshold could result in unwanted damage to internal components.

With smart grid technologies taunting new abilities such as energy management and real-time pricing, there has been much recent discussion on what distribution systems of the future can and should look like. It is important to understand the characteristics of the smart grid, in order to develop an approach that is robust enough to apply in today's environment as well as integrate seamlessly into the smart grid to take advantage of the opportunities it purports.

2.6 Smart Grid

The power utility industry has been utilizing advances in communication and IT over the years in order to improve efficiency, reliability, security and quality of service [47]. Momentum for the smart grid vision has increased recently due to policy and regulatory initiatives [17] [63].

The smart grid is envisioned to take advantage of all available modern technologies in transforming the current electrical grid into one that functions more intelligently. There are numerous potential benefits that smart grid technology is expected to facilitate, including:

- Displacement of about half of our nation's net oil imports [37]
- Better situational awareness and operator assistance [47]
- Reduction in U.S. carbon dioxide emissions by about 25 percent [37]
- Integration of renewable resources including solar, wind, and various types of energy storage [47]
- Reductions in emissions of urban air pollutants of 40 percent to 90 percent [37]

There are several newly developed and/or tailored smart grid attributes to help the nation achieve these advertised benefits.

2.6.1 Attributes. The first such attribute that must be understood is demand response. Demand response allows consumer load reduction in response to emergency and high-price conditions on the electricity grid [59]. Such conditions are more prevalent during peak load or congested operation.

The second relevant attribute of smart grid technology is its implementation of load rejection. Load rejection as an emergency resource to protect the grid from disruption is well understood and is implemented to operate either by system operator or through underfrequency and/or under-voltage relays [47]. The smart grid enhancement is that load rejection schemes can be enhanced to act more intelligently and be based on customer participation.

2.6.2 Threats and Vulnerabilities. Although threats of economic and industrial sabotage have long existed, the international proliferation of the Internet makes cyber economic and industrial sabotage an especially daunting and potentially economy-crippling threat [41] [65].

As we can see, vulnerabilities have not gone away (and maybe got worse). Therefore, it is important to understand how to recover from a situation once it happens. It will happen (cite instances).

Several other grid-related impacts are likely to emerge when adding a significant new load for charging plug-in hybrid vehicles. Higher system loading could impact the overall system reliability when the entire infrastructure is used near its maximum capability for long periods [37].

”Over the past several years, we have seen cyber attacks against critical infrastructures abroad, and many of our own infrastructures are as vulnerable as their foreign counterparts,” Director of National Intelligence Dennis Blair recently told lawmakers. ”A number of nations, including Russia and China, can disrupt elements of the U.S. information infrastructure [25]”. The growing reliance of utilities on

Internet-based communication has increased the vulnerability of control systems to spies and hackers, according to government reports.

It is not wise to imagine any internet-based communications to be completely secure and free from attack. The U.S. electrical grid is no different. Instead, researchers and developers must focus their resources and efforts to adapting to and overcoming such malicious intrusions and even routine equipment malfunctions. One such mechanism that is employed within the electric power systems is the Special Protection System (SPS) (also known as Special Protection Scheme).

2.7 Special Protection System

A special protection system is specifically designed to detect abnormal system conditions, preserve system stability and are designed to take pre-planned corrective action in response to certain disturbances, to mitigate the consequence of abnormal conditions [4][38][67]. These systems are often perceived as an attractive alternative to constructing new transmission lines because they can be placed in service relatively quickly and inexpensively.

In their most recent survey, CIGRÉ, the Council on Large Electric Systems, identified 113 special protection schemes in operation [66]. Additionally, the IEEE CIGRÉ survey concurred with these results and identified generator rejection as the most commonly used SPS [4]. The most common SPS types are consolidated in Table 2.4.

Related work in this field investigated the creation of a SPS that estimated load shedding levels under transient situations by using communication from regional generators and key loads [21][32][54]. This research will also focus on these aspects of an SPS, as generation rejection and load shedding are the most common responses employed by SPSs worldwide[3]. SPSs are designed to preserve system stability in the face of a large variety of disturbances, helping to prevent violent and disastrous effects.

Table 2.4: Percentages of Most Common SPS Types [3][66]

Types of SPS	Percentage
Generator Rejection	21.6
Load Rejection	10.8
Underfrequency Load Shedding	8.2
System Separation	6.3
Turbine Valve Control	6.3
Load & Generator Rejection	4.5
Stabilizers	4.5
HVDC Controls	3.6
Out-of-Step Relaying	2.7
Discrete Excitation Control	1.8
Dynamic Braking	1.8
Generator Runback	1.8
Var Compression	1.8
Combination of Schemes	11.7
Others	12.6

An example of such a disturbance is possible in systems that are interconnected by long or weak tie lines, which may be heavily loaded. When this scenario occurs, the power system may break apart in ways that are not predictable and possible create power system islands having large generation-to-load imbalances. Islanding occurs when a distributed generator (or group of distributed generators) continues to energize a portion of the utility system that has been separated from the main utility system [6]. It is not desirable for a distributed generator to island any part of the utility system as this can lead to safety and power quality problems such as the generation-to-load imbalance previously

addressed. Regardless of the specific implementation of an SPS, there are several traits in common that each will possess.

2.7.1 Traits. Protective schemes have at least four traits in common [4] that are pertinent to this research. First, all SPS implementations are dynamic security control systems and are designed to control power system stability in cases where the uncontrolled response is likely to be more damaging than the controlled response.

Secondly, all are devised by off-line analysis, as opposed to on-line real-time control. The reasons for this is that the power system response is too fast to allow for the usual sequential control system logic, which might be summarized as:

- make the observations in real time
- determine the scope of the disturbance
- decide what action is required, and then
- take the needed action

Third, many of these schemes are armed or disarmed, as required, in order to meet the needs of the system at a particular time. In other words, the special control logic may not be required under certain operating conditions, in which case the SPS is disarmed.

Finally, all of the schemes provide a particular type of remedial action that is designed to alleviate a certain observed system condition, or to take a predetermined action when a certain event occurs whose resulting effects are calculated to be too serious to ignore

Although the schemes have several traits in common, specific preplanned courses of action must be determined and tailored through detailed stability studies. Therefore, it is important to fully understand stability as it pertains to the power system.

2.7.2 Power System Stability. Power system stability may be broadly defined as that property of a power system that enables it to remain in a state of operating

equilibrium under normal operating conditions and to regain an acceptable state of equilibrium after being subjected to a disturbance. Instability within of a power system can be influenced by a wide range of factors and can take many different forms. Similarly, there are numerous events that can introduce instability into a power system. Typical event disturbances are identified below:

- Transmission faults
- Cascading outages of lines
- Generation outages
- Sudden, large load changes
- Combination of the above

There is much caution in the realm of power system stability, with numerous reminders that solutions to stability problems of one category should not be at the expense of another [39]. The basic operating requirements of an ac power system are that the synchronous generators must remain in synchronism and the voltages must be kept close to their rated values [51]. The capability of a power system to meet these requirements in the face of possible disturbances (line faults, generator and line outages, load switchings, et cetera (etc.)) is characterized by its transient, dynamic and voltage stability [55].

The simulations in this experiment have a great impact on the transient stability of the power grid and is what will be measured throughout each experiment. Transient stability is the ability of the power system to maintain synchronism when subjected to a severe transient disturbance [39]. Power system stability depends greatly on both the initial operating state and the severity of the disturbance. Although disturbances can vary greatly, as previously identified in Section 2.7.2, the SPS is designed and operated so as to be stable for a wide range of contingencies. Among the most common SPS types is load shedding and generator rejection as depicted in Table 2.4.

2.7.3 Generation Rejection Scheme. A Generator Rejection Scheme (GRS), when properly operating, significantly improves response following a contingency [69]. A GRS is designed to trip pre-selected generating unit(s) at a plant in order to prevent loss of the entire plant. Utilizing generation rejection to attempt to regain system stability is not a new approach. Quite to the contrary, generator rejection actually comprises the most common type of special protection scheme, as identified in Table 2.4. The selective tripping of generating units for severe transmission system contingencies has been used as a method of improving system stability for many years [39]. The approach of generator tripping as a stability aid was initially confined to hydro plants, but has gradually extended to fossil-fuel-fired and nuclear units since the *1970s*. Even with generation rejection implemented, the creation of a load imbalance in a power system may cause such an excess of load over generation that there is no alternative but to shed some of the load.

2.7.4 Underfrequency Load Shedding. In many cases, underfrequency conditions arise due to the breakup of a large system into two or more islands. It is often necessary to install load shedding relays throughout the power system so that any possible island configuration will be protected against underfrequency operation.

An important aspect of load shedding is that it is necessary for all of the utilities that make up the interconnected system to come to an agreement as to the amount and timing of load shedding, so that all portions of the system behave in approximately the same manner when load shedding is required, irrespective of the exact cut set that defines the separation [3].

A desirable and obvious, yet not trivial, underlying requirement of load shedding is the fact that a load should shed when called upon to do so. It is desirable to improve the operation of load shedding by selecting only those loads that the system could trust to do what is instructed of them in a time of need. Researchers are currently approaching this facet of load shedding through the use of trust management [21][32][54].

2.7.5 Trust Management. Although the idea of trust relates to firm beliefs in attributes such as reliability, honesty and competence, it has proven a difficult topic to research. This is primarily due to a lack of consensus in literature on the definition of trust and exactly what constitutes trust management [2] [27] [46].

The definition of trust for the purposes of this research is [5]:

Trust is firm belief in the competence of an entity to act as expected, such that this firm belief is not a fixed value associated with the entity, but rather it is subject to the entity's behavior and applies only within a specific context at a given time.

With regard to the power grid, the concept of trust management or a trust system is to provide software agents that plug into an existing network, somewhat transparently, to perform the functions of correlating data and identifying risk levels for corresponding events and status updates to point to negative impacts on utility services. Researchers have developed such trust systems that operate by intercepting messages or commands from network nodes and validates input to identify security risks or bad data [15].

2.7.6 Reputation Based Trust. Most research on reputation based trust utilizes information such as community-based feedbacks about past experiences of peers to help make recommendations and judgements on quality and reliability of the applicable transactions [68]. A challenge with any reputation based trust system is how to deal with the malicious behaviors of peers, or malfunction that presents itself as such. For the purpose of this research, a node displaying malicious or malfunctioning behavior is identified as *bad*.

Reputation based trust is a topic of much interest and research in recent years [2] [9] [44], and each presents a manner to cope with such *bad* participants. The research in this thesis, however, focuses specifically on the electric power system and addressing challenges specific to it.

Therefore, there are several researchers of keen interest, as they have effectively demonstrated the application of reputation based trust within electrical power systems [10] [20] [54]. Regardless of the specific SPS being utilized, taking correct action at the correct time is key to success. This research determines the value of applying a concept to an SPS that has not been yet been documented; namely modifying traditional exponentially weighted moving average calculations. There is a push for modernization, which is an ongoing process. While planning for the future, the effort must continue to add security and reliability into the existing SCADA electric utility equipment. Adding a reputation based trust system that can be optimized for each application not only adds reliability now, but also provides a vehicle to add new enhancements directly into the smart grid as it grows and evolves.

2.7.7 Special Protection System Trust Module. The specific trust implementation in this research is an adaptation of existing research and has three major components; a trust assignment component, a fault detection component and a decision component [21]:

1. The trust assignment component uses context sensitive information and periodic intercommunication messages to determine individual smart grid protection components' trust values. The context sensitive information shared by smart grid protection components are generator frequencies, which is used to reach a consensus concerning the state of the system. Protection components in agreement with the consensus are assigned a high trust value and those whom disagree are assigned a low trust value.
2. The fault detection component monitors one or more predetermined values for changes that indicate a condition requiring corrective action, namely grid frequencies measured at all generator and load locations.

The decision component is notified when a monitored value exceeds or falls below it predetermined limits.

3. The decision component uses the previously assigned trust values to validate the detected fault and responds appropriately to minimize power grid downtime.

To dynamically determine the appropriate corrective action, the referenced trust module utilizes a greedy algorithm approach.

2.7.8 Greedy Sorting Algorithm. The greedy sorting algorithm is used by the trust management module to determine the order in which the protection system nodes are selected for load shedding, by using assigned trust values, node type and load values. In this manner, the trust module sorts all protection system nodes. Table 2.5 is an example set of sorted nodes with precedence from left to right, i.e., first sort by *Type* (type of node) and then by *Trust Value* (nodes calculated trust value) followed by *Load in Megawatts (MW)* (load amount at the node, customer authorized load shed amount (20% of load amount in this example) and the node's Identification Number (ID) [21].

Table 2.5: Sorted Nodes For Possible Load Shedding [21]

Type	Trust Value	Load (MW)	Shed Amt (MW)	Node ID
Load	High	1812	362	120
Load	High	1492	298	73
Load	High	1492	298	25
Load	High	1250	250	72
Load	Low	1590	318	33
Load	Low	1492	298	82

If a frequency disturbance is detected and requires the power grid to shed 700 MW of power, then the greedy algorithm would attempt to meet this requirement by selecting the first load node in Table 2.5, namely node 120 [21]. Since the selection of this node is not

enough to meet the 700 MW of power required to shed, the greedy algorithm selects the next node, in this case node 73. This process must continue one more round to include node 25, ensuring that the shed amount is great than or equal to 700 MW. Now the selected node have enough load available for shedding (i.e., 958 MW) to meet the requirement—enabling the greedy algorithm to stop selecting additional nodes for load shedding [20].

It is important that if the greedy algorithm exhausts all load nodes from Table 2.5 with a trust value identified as *High* before it reaches the required load shed amount, it then selects *Low* trust values until the required amount is met. Since the untrusted nodes are not expected to follow the shed request, the system often fails to maintain the required frequency threshold.

A second topic of interest is the specific determination of *High* and/or *Low* trust values. Related research determines this binary value in a number of different manners [10] [21] [54]. One option is to simply apply a trust threshold to the final observation. For example, if the final frequency observed before the trust determination is made is within tolerance, then the node would be trusted. If not within tolerance, then the corresponding trust value for that time step would be *Low*, and would be the nodes overall trust value if the trust determination is made during that time step. This aligns with the notion of a traditional EWMA implementation with a lambda (λ) = 1, as discussed in Section 2.8.3.

Other trust management researchers have included a history and chose to equally weight each of the trust value observations made for a specified period of time. This concept equally weighs all past observations and corresponds to a traditional EWMA implementation with a $\lambda = 0$, also discussed in Section 2.8.3.

2.8 Exponentially Weighted Moving Averages

Several approaches have been taken to determine the best application of trust [10] [21] [54]. These approaches implement a variety of the SPS protection tools available in

Table 2.4. It is the asserataion of this research that an EWMA is applicable to help determine the trust value within a reputation based trust SPS implementation.

2.8.1 Overview. The EWMA concepts were first introduced in 1959 [53]. Although the EWMA is known to have optimal properties in some forecasting and control applications [12] [48], it has largely been neglected as a tool by quality-control analysts [43].

An exponentially weighted moving average is a means of smoothing random fluctuations that has the following desirable properties [30]:

1. Declining weight is put on older data
2. It is extremely easy to compute
3. Minimum data is required

Observations are assumed to be sequentially recorded and these observations, or some functions thereof, are usually plotted for the purpose of controlling a manufacturing process. Additionally, the desire to employ historical data more resourcefully has occasionally led to the notion of the moving average [33]. For example, a plot of a moving average of $k = 8$ observations will only display the average of the eight most recent observations. This is a first-in-first-out implementation, where newer data forces the older data out of the computation.

2.8.2 Sample Calculations. Traditional EWMA implementations give less and less weight to data as they get older and older. A new value is easily obtained by computing a weighted average of two variables, namely the value of the average of the previous period and the current value of the variable.

A simply application of an EWMA is presented in as Equation 2.1 [30], which follows the rule: take a weighted average of *all* past observations and use this as your

forecast of the present mean of the distribution.

$$\bar{S}_t = B[S_t + AS_{t-1} + A^2S_{t-2} + A^3S_{t-3} + A^4S_{t-4} + \dots] \quad (2.1)$$

where B is a constant between 0 and 1, A is $(1 - B)$, the S s are observations of the variable and the t subscript indicates the time ordering of the observations. \bar{S}_t is the estimate of the expected value of the distribution.

2.8.3 Traditional EWMA Example. This thesis research is based upon a variation of the traditional application of EWMA [43], based on the statistic identified in Equation 2.2, which presents the symbol Z_i .

$$Z_i = \lambda Y_i + (1 - \lambda)Z_{i-1}, \quad 0 < \lambda \leq 1, \quad i = 1, 2, \dots, n \quad (2.2)$$

where

- Z_0 is the mean of the historical data
- Y_i is the observation at time i
- N is the number of n observations to be monitored

The starting value Z_0 is often realized as the target value of the monitored process [56]. This approach is necessary there is no target value of the process being monitored. Utilize Equation 2.3 to simply determine the average:

$$Z_0 = \bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i, \quad i = 1, 2, \dots, n \quad (2.3)$$

Additionally, the sequentially recorded observations, Y_i , are individually observed values from the process.

TRADITIONAL EWMA EXAMPLE:

The given a data set, Y , has 8 observations, so $N = 8$:

$$Y = \{1, 1, 1, 1, 1, 1, 0, 0\}.$$

Next calculate Z_0 , which in this case is equal to \bar{Y} ,

$$\text{Which is } Z_0 = \bar{Y} = \frac{1}{8}\{1, 1, 1, 1, 1, 1, 0, 0\} = 0.75.$$

Implementing Equation 2.2 with $\lambda = 0.6$ yields the following initial Z_i values:

$$Z_0 = 0.75$$

$$Z_1 = \lambda Y_1 + (1 - \lambda)Z_0 = 0.6 \times 1 + (1 - 0.4) \times 0.75 = 0.9000$$

$$Z_2 = \lambda Y_2 + (1 - \lambda)Z_1 = 0.6 \times 1 + (1 - 0.4) \times 0.9000 = 0.9600$$

Following the same rules, all the Z_n values are presented in Table 2.6.

Table 2.6: Intermediate Calculations of a Traditional EWMA Algorithm [43]

Z_0	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8
0.75	0.9000	0.9600	0.9840	0.9936	0.9974	0.9990	0.3996	0.1598

Figure 3.1 captures the resulting EWMA chart for data set Y . It is important to note that although only the last two observations from Y were 0's, the final EWMA is 0.1598, which is due to the fact a traditional EWMA property is that declining weight is placed older data, given $0 < \lambda \leq 1$.

To better understand the role λ has in the creation of Figure 3.1, where the final calculated average is 0.1598, consider the cases where λ is equal to 0 and the case where it is equal to 1. $\lambda = 0$ is essentially giving equal weight to all observations, which is equal to the average, of 0.75 in this case. Conversely, $\lambda = 1$ only gives weight to the most recent observation, yielding a calculated average of 0, as this is the final value in data set Y .

Therefore, given data set Y , the resultant value given a λ of 0, 0.6 and 1.0 yields calculated averages of 0, 0.1598 and 0.75 respectively. If applied to a trust scheme, data set Y may be trusted on some occasions and not trusted on others, depending on the value of λ and also the trust threshold. Selecting appropriate λ and trust thresholds are key to successful implementation of an EWMA scheme within a reputation based trust management system.

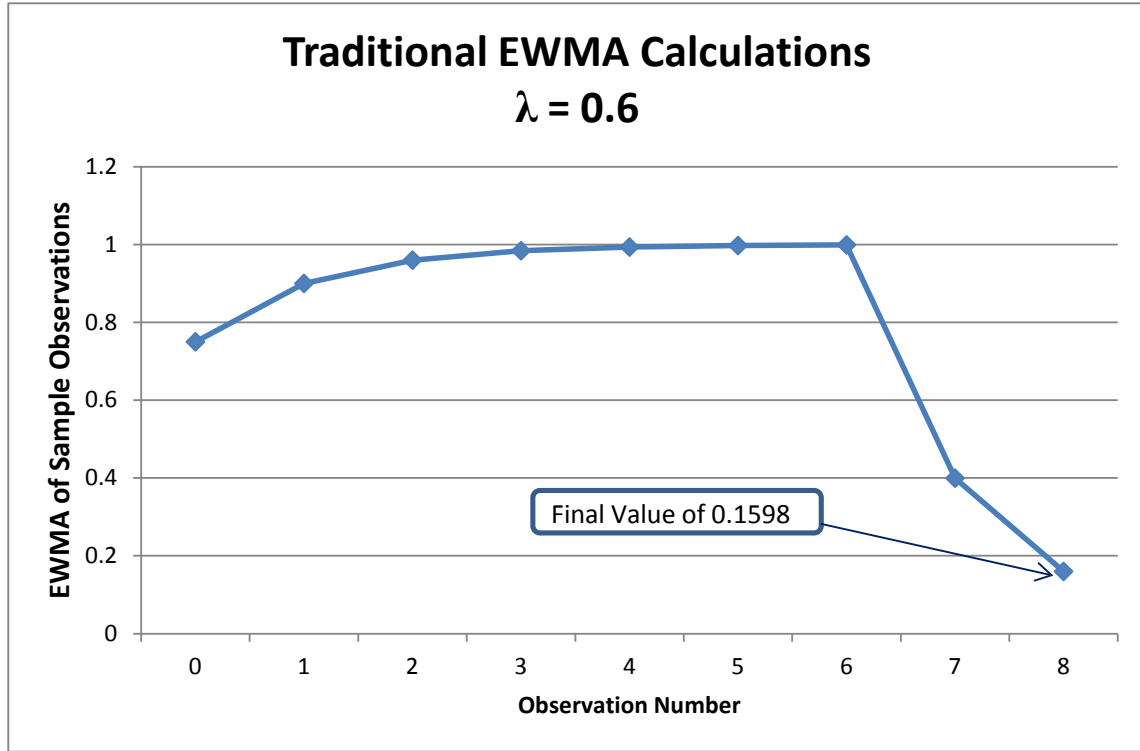


Figure 2.3: Traditional EWMA Graph on Data Set Y , with $\lambda = 0.6$

2.8.4 SPS Applicability. In the case of an SPS incorporating reputation based trust, there is no target value for Z_0 . This is due to the fact that a particular node may be either good or bad (malicious or malfunctioning). Therefore, although the minimum acceptable frequency for normal operation may be 58.8 Hz, initializing Z_0 to this value could create a situation where bad nodes are not detected.

The overall goal of the SPS is like many industrial processes, where the goal is to maintain a stable state or in this case frequency. This is not the case, however, for the reputation based trust management implementation, where the goal is to accurately reflect trustworthiness of each node regardless of whether it is behaving good or bad.

Lastly, due to the complexities associated with SPS implementation, the traditional EWMA implementation cannot be directly applied. In this thesis, several changes are made to the algorithm described by Equation 2.2 before it is applied to any trust

management. Specific changes to the algorithm along with detailed testing methodology is presented in Chapter 3.

3 Methodology

3.1 Overview

Thorough and rigorous testing is important to validate assumptions and determine causality. The goal of creating the simulation environment in this thesis is not to necessarily optimize individual run-time environment components. Rather, the goal for utilizing this particular experimental environment is to create an environment that is conducive to robust reputation based trust testing. Finally, the role of statistics in the scientific method cannot be overstated.

The discipline of "statistics" can be described as the art and science of using quantitative information (data) to gain understanding and to make informed decisions [40]. Therefore, the methodology in this thesis was designed to facilitate the condensing of large volumes of data into forms that facilitate understanding.

Special protection systems are evaluated in terms of its ability to take correct actions during disruptions within electrical systems such as those identified in Table 2.3. The overall goal of these protective systems is quite simply to preserve system stability.

This research methodology explores the applicability of a proposed special protection system that calculates reputation based trust based upon a modified Exponentially Weighted Moving Averages (EWMA) algorithm developed within this research. To accomplish this, the research methodology is divided into stages that facilitate data collection and analysis.

3.2 Modified Exponentially Weighted Moving Average

Due to the properties of a special protection scheme, namely that it is called after a catastrophic failure is observed, direct application of the traditional exponentially weighted moving average equation, Equation 2.2, is not possible. This research assumes

that during normal operation, applicable system measurements and their resultant trust values are consistent and accurate. Then when a special protection condition arises, the stability of the system is jeopardized, as discussed in Section 2.7.2. Therefore, the traditional EWMA property that *declining weight is put on older data* simply will not suffice.

This research explores the application of reputation based trust that implements a modified EWMA scheme. If implemented correctly, the protection system will correctly react to the disturbance before the system reaches the minimum frequency threshold of 58.8 Hz, while minimizing both false positives and false negatives.

Due to the instability induced by the event requiring the special protection system action, *older data should have more weight* than the newer data collected during periods of instability. Therefore, critical modifications must be made to the traditional EWMA algorithm to accurately account for the properties surrounding SPS implementation.

Equation 2.2 identified the traditional EWMA algorithm:

$$Z_i = \lambda Y_i + (1 - \lambda)Z_{i-1}, \quad 0 < \lambda \leq 1, \quad i = 1, 2, \dots, n$$

Within this research, consensus information as identified in Section 2.7.8, is treated as intermediary trust values. Therefore, instead of making trust decisions on data set Y directly, as occurred in the traditional EWMA calculations in Section 2.8.2, data set Y is now viewed as a set of a intermediary trust value. The modified algorithm is applied to data set Y to determine the actual trust values.

When applied to data set $Y = \{1, 1, 1, 1, 1, 1, 0, 0\}$ with a $\lambda = 0.6$, the resultant calculation, or trust value in this case is 0.1598 as identified in Table 2.6. This research attests that the final two 0's in data set Y are not indicative of an untrustworthy condition, but rather should be expected during an SPS condition.

To correct for this SPS property, it is important to traverse the data set backward, or from most recent to oldest and give the most weight to older data. To accomplish this

result, Equations 2.3 and 2.2 must be modified to become Equations 3.1 and 3.2 respectively:

$$\hat{Z}_{N+1} = \bar{Y} = \frac{1}{N} \sum_{j=1}^N Y_j, \quad (3.1)$$

$$\hat{Z}_j = \lambda Y_j + (1 - \lambda)\hat{Z}_{j+1}, \quad 0 < \lambda \leq 1, \quad j = n, n-1, \dots, 1 \quad (3.2)$$

where

- \hat{Z}_{N+1} is the mean of the historical data
- Y_j is the observation at time j
- N is the number of n observations to be monitored
- \hat{Z}_j is the modified EWMA algorithm

These new equations allow for revised trust calculations from those realized in Section 2.8.3.

MODIFIED EWMA EXAMPLE:

Once again, the given a data set, Y , has 8 observations, so $N = 8$:

$$Y = \{1, 1, 1, 1, 1, 1, 0, 0\}.$$

Next calculate \hat{Z}_{N+1} , which is equal to $\bar{Y} = \hat{Z}_9$, since $N = 8$.

$$\text{Which is } \hat{Z}_{N+1} = \hat{Z}_9 = \bar{Y} = \frac{1}{8}\{1, 1, 1, 1, 1, 1, 0, 0\} = 0.75.$$

Implementing Equation 3.2 with $\lambda = 0.6$ yields the following initial \hat{Z}_j values:

$$\hat{Z}_9 = 0.75$$

$$\hat{Z}_8 = \lambda Y_8 + (1 - \lambda)\hat{Z}_9 = 0.6 \times 0 + (1 - 0.4) \times 0.75 = 0.3000$$

$$\hat{Z}_7 = \lambda Y_7 + (1 - \lambda)\hat{Z}_8 = 0.6 \times 0 + (1 - 0.4) \times 0.3000 = 0.1200$$

$$\hat{Z}_6 = \lambda Y_6 + (1 - \lambda)\hat{Z}_7 = 0.6 \times 1 + (1 - 0.4) \times 0.1200 = 0.6480$$

Following the same rules, all the \hat{Z}_j values are presented in Table 3.1.

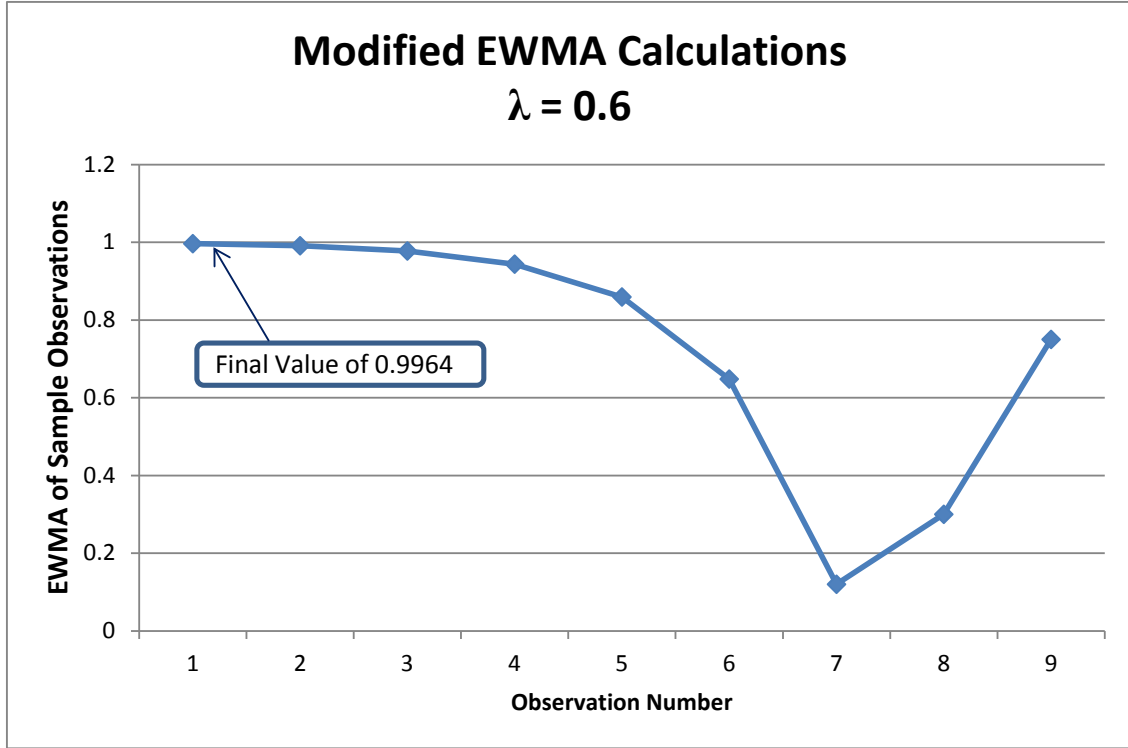


Figure 3.1: Modified EWMA Graph on Data Set Y , with $\lambda = 0.6$

Table 3.1: Intermediate Calculations of Modified EWMA Algorithm

\hat{Z}_1	\hat{Z}_2	\hat{Z}_3	\hat{Z}_4	\hat{Z}_5	\hat{Z}_6	\hat{Z}_7	\hat{Z}_8	\hat{Z}_9
0.9964	0.9910	0.9775	0.9437	0.8592	0.6480	0.1200	0.3000	0.7500

Figure 3.1 captures the resulting EWMA chart for data set Y . It is important to note that this revised algorithm generates a markedly different result than the traditional implementation in Section 2.8.3, resulting in a final EWMA (trust calculation) of 0.9964. Since declining is placed on more recent data, the final two values of data set Y , which are both 0's, the resulting calculation is quite high. This final value would be considered trusted given any threshold trust below 0.9964. Critical to this research is implementation

of this modified exponentially weight moving averages algorithm to determine optimal λ and trust threshold values, as determined using techniques in Section 3.2.1.

3.2.1 Analysis. Receiver operating characteristics graphs have long been used in signal detection theory to depict the tradeoff between hit rates and false alarm rates of classifiers, and have since applied to machine learning in the evaluation and comparison of algorithms [19] [62].

		Observed	
		TRUE	FALSE
Predicted	TRUE	True Postiive (TP)	False Positive (FP)
	FALSE	False Negative (FN)	True Negative (TN)

Figure 3.2: Confusion Martix [22]

Labeled in Figure 3.2 are the four possible outcomes of a given classifier and instance. If the instance is positive and classified as positive, it is counted as a True Positive (TP); if it is classified as negative, it is counted as a False Negative (FN). Similarly, if the instance is negative and it is classified as negative, it is counted as a True Negative (TN); if it is classified as positive, it is counted as a False Positive (FP). Given a classifier and a set of instances, (the test set), a two-by-two *confusion matrix* can be constructed representing the dispositions of the set of instances [22].

Two additional performance metrics must be identified to construct ROC curves, namely the True Positive Rate (TPR) and False Positive Rate (FPR) as identified in Equations 3.3 and 3.4 [29].

$$\text{TruePositiveRate (TPR)} = \frac{TP}{TP + FN} \quad (3.3)$$

$$\text{FalsePositiveRate (FPR)} = \frac{FP}{FP + TN} \quad (3.4)$$

ROC curves are comprised of single points comprised of (FP rate, TP rate) pairs, known as discrete classifiers [22], as depicted in Figure 3.3.

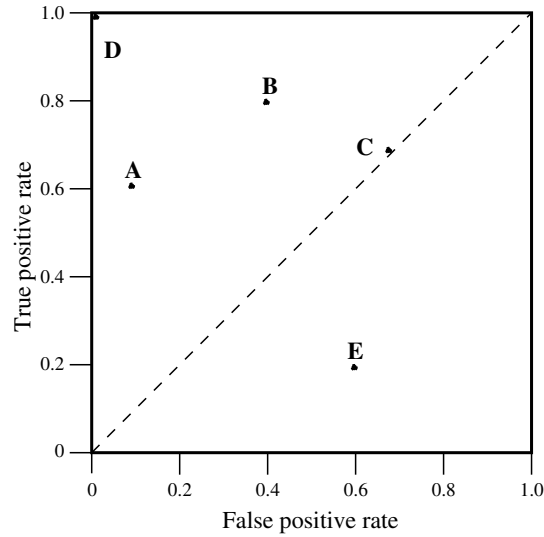


Figure 3.3: Basic ROC Graph Showing Five Discrete Classifiers [22]

The lower left point (0,0) represents a strategy that commits no false positives, but also no true positives. The opposite strategy would be depicted by a point in the upper right (1,1). Unfortunately, this strategy unconditionally issues both true and false positives. Informally, one point in ROC space is better than another if it is to the northwest (TP rate is higher, FP rate is lower, or both) of the first [22]. Finally, the diagonal line $y = x$

represents a strategy of randomly guessing a class. With regard to the points plotted in Figure 3.3, D would be the optimal strategy of the discrete classifiers available, as it serves to both maximize true positives and minimize false positives. Research in this thesis will consider a classifier optimal if it accomplishes both of these objectives as well.

3.3 Testing Environment

The research contained in this thesis makes use primarily of the EPOCHS Simulation Environment, PSS/E electromechanical transient simulator and NS2 network simulator.

3.3.1 EPOCHS Simulation Environment. EPOCHS is a simulation platform that integrates multiple research and Commercial Off-The-Shelf (COTS) systems to bridge the gap [32]. It allows users to investigate electromechanical scenarios using PSS/E and NS2. The focus of EPOCHS is to integrate power and network communication simulators so that their internal simulation time clocks advance seamlessly.

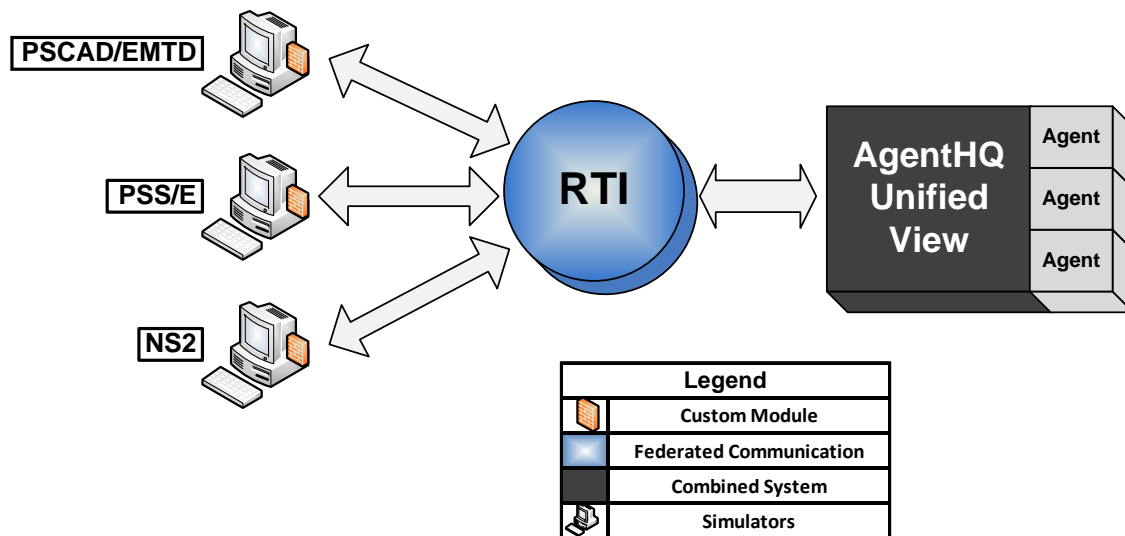


Figure 3.4: Relationship Between EPOCHS Components

3.3.2 *Network Simulator 2.* Network Simulator (Version 2), widely known as NS2, is an event-driven simulation tool that is useful in studying the dynamic nature of communication networks and has gained popularity in the networking research community since its birth in 1989 [35]. As identified in 3.5, NS2 takes in the name of a Tcl simulation scripting file as its argument.

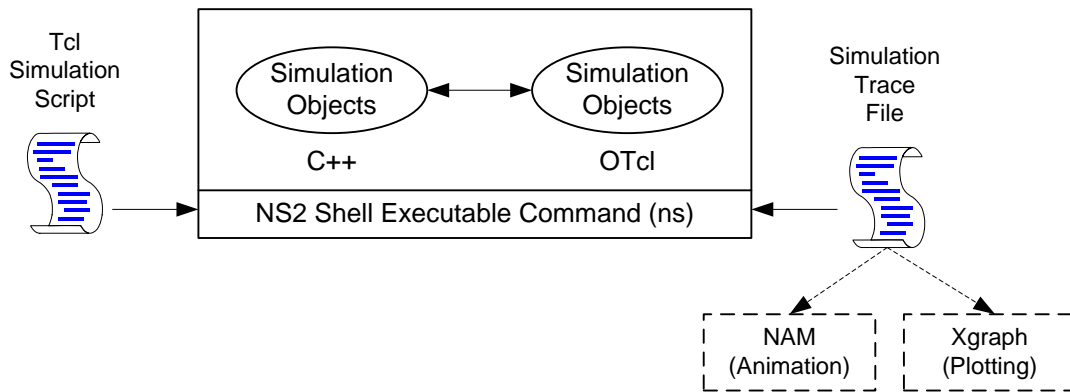


Figure 3.5: NS2 Architecture [35]

Figure 3.6 provides a representation of smart grid implementation that NS2 attempts to accurately model. Specifically, there are several different types of nodes that are interconnected such as customers, substations power plants and control centers. Within NS2 each of these node types is represented by a software agent, where software agents are autonomous software entities designed to mimic the behavior of real world systems, and each would be strategically located within Intelligent Electronic Devices (IEDs) in a real world environment [21].

Within NS2, the software agents communicate with each other, share data and make observations of peers.

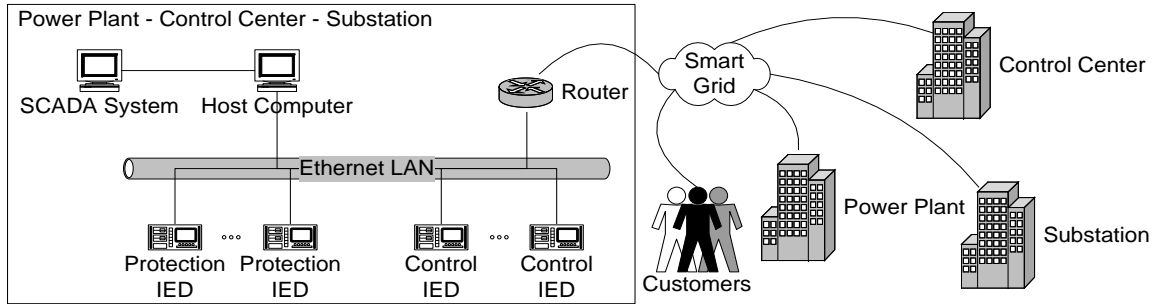


Figure 3.6: Representation of a Smart Grid Wide Area Network [32]

3.3.3 PSS/E. PSS/E is the premier software tool used by electrical transmission participants world-wide. Since its inception in 1976, it has become the most comprehensive, technically advanced, and widely used commercial program of its type [60]. Software agents communicate with PSS/E through the EPOCHS environment, as depicted in Figure 3.4. As such, each software agent can access and modify their corresponding power component's data (e.g., access sensor data, engage relays, change load power levels, etc.), ensuring seamless integration of proposed the trust management enabled exponentially weighted moving averages algorithm capability into a simulated smart grid enhanced power grid.

3.3.4 AgentHQ. AgentHQ presents a unified environment to agents and acts as a proxy when agents interact with other EPOCHS components.

In order to support the operation of software agents on the power grid, a hardware device is needed that has the computational, communication and I/O capabilities to meet the agent demands. EPOCHS uses agent-based intelligent electronic devices for this purpose so software agents can perform the necessary protection and control functions needed [31].

Through AgentHQ, agents can get and set messages to each other.

3.3.5 RTI. The "glue" holding the high level architecture combinations together is a control component known as a Runtime Infrastructure (RTI). The RTI routes all messages between simulation components and ensures that the simulation time is synchronized across all components [32]. EPOCHS implements a time-stepped model, in which each of the component federates executes until a preset simulation time is reached. In this simulation model, the amount of time between synchronization points has a fixed length.

3.3.6 Component Interaction. Synchronization of simulators within EPOCHS follows a simple algorithm. As soon as NS2 and PSS/E begin execution, the RTI halts the simulators and waits for synchronization messages from both the power simulator and NS2. The RTI then yields control to AgentHQ, who in turn passes control on to the agents one by one until all have executed.

During the simulations, this is where agents are sending communication messages and getting/setting power system variables, which is the basis of how the underlying trust calculations are made. Once all agents have executed, AgentHQ returns control back to the RTI, who in turn notifies both NS2 and PSS/E that the current time step is done. The two simulation engines run for at least one additional time step, ensuring that no more actions are required.

3.4 System Studied

Experimental simulations make use of a modified version of IEEE's 145-bus 50-generator test case [64]. The system has been modified by adding a 500 kV line from bus 1 to bus 25. The rational of this additional line is to create a situation that requires the use of a special protection system in order to maintain stability.

In general, power systems can sustain the loss of a single tie line. However, most power systems require remedial action with the loss of a second line, if the line is not

cleared quickly enough. An additional modification to the IEEE test case is that total system capacity has been reduced. The lower system capacity makes the power flow along the main corridor much more important than it is in the original system.

In the IEEE test case, the main SPS agent is located at bus 1, and it identifies extreme contingencies such as the loss of two lines. It then performs both generation rejection with preset units and load shedding based upon real-time measurements [64]. The specific generator to be rejected was determined through simulation studies [32].

Utilizing this modified IEEE 145-bus 50-generator test case, each simulation implements the same basic template:

1. Power transmission lines are tripped due to malfunction and overloading
The result is a power transmission system imbalance
2. Preplanned action requires removal of a specific generator (93) from the system
The result is an imbalanced and unstable system frequency drop
3. The special protection system relies on calculated trust values of each load node to intelligently shed load

The goal is regaining system stability by shedding the required amount of load

The specific scenario for this research has been created to allow for robust trust measurement experimentation.

3.4.1 Special Protection Scheme Action Goal. A key indicator of success throughout this research is the ability of the system to hold the system's frequency above 58.8 Hz as illustrated in Figure 2.2. The special protection scheme implemented in this research is based upon the original EPOCHs research, which employs an algorithm to determine the precise generation shortfall when a disturbance occurs [32]. Specific remedial SPS actions are dependant upon the system in which they are applied.

3.5 Test Scenario

The overall goal of an SPS is to prevent instability and preserve the system's integrity within a safe operating range or quickly recover from a critical condition. Critical disturbances are those expected to have a devastating effect on the power system under a particular operating condition [3].

The modified trust management SPS test scenarios monitor the power grid's frequency for disturbances that are indicative of an imminent fault and attempts to prevent the fault by generation rejection (Section 2.7.3 and load shedding Section 2.7.4) [21]. In this test case, two inter-tie power lines are lost causing a protection system condition, causing the power grid to become transiently unstable, necessitating power generation rejection.

During simulation on the system described in Section 3.4, dropping generator 93 constitutes a form of dynamic security assurance, where the actions to be taken in response to a given condition are preplanned [3]. Since generating units can be rapidly tripped, this is a very effective and efficient means of improving transient stability. For this scenario, it has been determined through stability studies, that a serious enough condition exists to call for preplanned control action.

The main SPS agent communicates with generation and load agents to gather data values and also communicates with agents located at major system or load buses to collect voltage and frequency measurements as well as load available for shedding [32]. Upon detection of the impending instability, it is necessary to shed selected loads so that transient stability is maintained without resorting to system separation, isolation of the affected region. The SPS employs an algorithm to determine the appropriate amount of load to shed in order to hold the system's frequency above 58.8Hz . The load agents are mainly located at distribution substations and shed load when ordered to do so by the main SPS agent. Specifically, within each simulation [32]:

A fault occurs on the line from bus 1 to bus 25 at time 0. The fault is cleared at $0.07s$ and a trip command is sent to generator 93 at $0.10s$. Since the fault is cleared after the critical fault-clearing time, the system becomes transiently unstable and one group of 17 generators loses synchronism with another group of 33 generators. The main SPS agent at bus 1 recognizes the situation and begins to communicate with other system agents to gather various data values and also sends a generation rejection order to agent at bus 93.

Generation rejection keeps the system stable, but without any corrective action, the frequency drops below the $58.8Hz$ threshold.

The main agent detects the "disturbance" created by generation rejection. It then estimates the disturbance size $1862MW$, calculates that there is $2090MW$ generation remaining and predicts that the steady-state frequency after the disturbance will be $57.45Hz$.

Therefore, although the rejection of generator 93 counteracts the power grid's transient instability, it causes an unacceptable decrease in frequency, where the supplied generator power is less than the load power demanded. Such a frequency drop could induce a blackout in the power grid similar to the 1965 northeast blackout [61]. Since the predicted steady-state system frequency of $57.45Hz$ is below the preset minimum frequency of $58.8Hz$, load shedding is required.

This mandatory load shedding is levied on selected load nodes. If the selected load shedding nodes are untrusted and refuse to load shed their fair amount, then the special protection system will fail to maintain the power grid's frequency above $58.8 Hz$ [32]. If the selected load shedding nodes are untrusted and refuse to load shed their fair amount, then the special protection system will fail to maintain the power grid's frequency above $58.8 Hz$ [32]. This underlines the importance of the trust management system's decision

module selection on trusted nodes for load shedding, recalling that trusted nodes have a higher probability of successfully completing assigned tasks than untrusted nodes [20].

The trust management's decision module selects nodes for load shedding based upon assigned trust values and the amount of load that must be shed. In this research, frequency information provided by all the nodes is used to determine individual trust values. The trust management decision module uses this information, and calculates modified exponentially weighted moving averages algorithm, to select nodes for load shedding and determine how much each selected node must shed [21].

The trust management decision module then sends each selected node a load shed message with the load shed amount required by the node, and the trusted nodes load shed their assigned amount, maintaining the power grid frequency above 58.8 Hz [20].

3.6 Experimental Design or Test Cases

Computer simulations are used to demonstrate the utility of modified trust management modules in special protection systems within a power grid. The protection systems are augmented as follows [21]:

1. The traditional special protection system is augmented with a trust management module
2. The assignment module utilizes current grid frequency information in a modified EWMA reputation based manner to establish and assign trust values
3. The fault detection module uses the traditional frequency disturbance mechanism to detect system conditions indicative of an imminent under-frequency fault
4. The decision module uses a greedy algorithm approach to determine which buses to select for load shedding

This normality testing is also conducted in related research approaches [10] [21] [54], and in this research constitutes Stage 1.

3.6.1 Stage 1 (Normality Testing). NS2 has 64 predefined good random seed values in their random number generator for computer simulation experiments, which are

equally spaced around a 2^{31} cycle of random numbers [35]. 36 seeds are chosen at random from NS2's predefined good random seed values. For Stages 1 and 3, these same 36 seeds (a set) are utilized so that data can be compared, attempting to minimize any introduced bias or unwanted variability.

The seeds in each set are used generate a listing of bad nodes that will not be known to the system during simulation. It is the job of the trust mechanism to effectively determine which nodes are good and bad. Normality testing is performed for each of the 5, 10 and 15 bad node test cases. Each of these are validated to ensure that the data generated are normal, allowing additional statistical analysis and inference.

Validation of Normality includes visual inspections of a Histogram and Quantile-Quantile (Q-Q) plots, along with further confirmation by the Shapiro-Wilk test [57]. Results that indicate the simulation generates Normal results means that research can continue to Stage 2.

3.6.2 Stage 2 (Modified EWMA). Determination of the optimal strategy is Stage 2 of this research and is based on the Receiver Operating Characteristic (ROC) principles identified in Section 3.2.1. Validation of the optimal strategy is based upon numerical validation of the True Positive Rate vs. False Positive Rate of the selected strategy as well as through graphical representation of all strategies on a ROC curve.

Stage 3 incorporates the optimal strategy into the trust module, identified in Section 2.7.7, and determines viability of this new approach.

3.6.3 Stage 3 (5, 10 and 15 Bad Node Frequencies). The goal of Stage 3 experiments is to determine the ability of the modified exponentially weighted moving averages algorithm, as compared to the special protection system without any trust management implemented, and related research that examines only the final trust value as a basis for trustworthiness (equal to $\lambda = 1$). The modified EWMA is tested using the

optimal λ and Trust Threshold values as determined in Section 4.3. The experiments are completed utilizing modified EWMA Equations, 3.1 and 3.2.

Measurements will focus on the frequency that each simulation is able to maintain, with the goal being the minimum acceptable value of 58.8 Hz. An Analysis of Variance (ANOVA) and comparison of Confidence Intervals (CI) are used to determine the statistical significance of the simulation results.

3.7 Analysis

Analysis techniques in this research range from Receiver Operating Characteristics (ROC) curves to determine appropriate λ and Trust Threshold values, Analysis of Variance (ANOVA) analysis, Quantile-Quantile (Q-Q) plot comparisons and a comparison of resulting confidence intervals.

3.8 Methodology Summary

A clear research methodology is essential to evaluate the hypothesis contained herein. In this research, the power transmission system and distributed special protection system are the component under test. Simulation is selected as the appropriate evaluation technique and the experimental design is identified to achieve a 99% confidence interval. Finally, this research methodology serves to identify a method to collect valid data required to evaluate and analyze the performance of the proposed modified exponentially weighted moving averages enabled trust system.

4 Analysis and Results

4.1 Overview

This section describes the results obtained from applying the design described in Chapter 3, and associated analysis. This analysis is both observational (e.g., identifying trends, showing success or failure, observing details of what the results show) and interpretive analysis (e.g., describing why results are the way they are, what underlying principles contributed to the success or failure).

4.2 Stage 1 (Normality Testing)

Stage 1 simulations were conducted in accordance with the methodology outlined in Section 3.6.1:

36 seeds are chosen at random from NS2's predefined good random seed values. For Stages 1 and 3, these same 36 seeds (a set) are utilized so that data can be compared, attempting to minimize any introduced bias or unwanted variability.

The seeds in each set are used generate a listing of bad nodes that will not be known to the system during simulation. It is the job of the trust mechanism to effectively determine which nodes are good and bad. Normality testing is performed for each of the 5, 10 and 15 bad node test cases. Each of these are validated to ensure that the data generated are normal, allowing additional statistical analysis and inference.

Represented here are the results from the 5 bad node test cases. If a determination of Normality can be confirmed, then one can infer that the random bad nodes are indeed chosen at random and that the data produced from simulation is representative of real-world results. To that end, outputs from this stage include the visual Normality

validation tools of a Histogram (see Figure 4.1) and Quantile-Quantile (Q-Q) plot (see Figure 4.2).

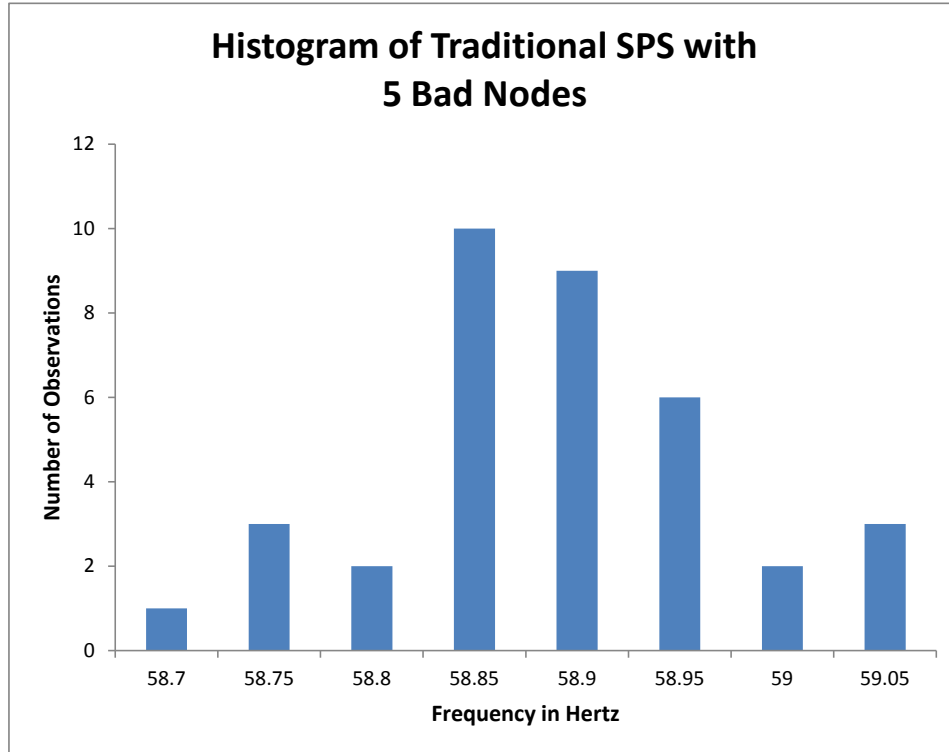


Figure 4.1: Histogram Generated from Simulation Results for Traditional SPS with 5 Bad Nodes)

The Histogram displays characteristics of Normality; namely, a resemblance to the desired bell curve. Additionally, the linearity of the points identified on the Q-Q plot suggest that the data are Normally distributed as well.

These positive results help support the notion of Normality and are cause to conduct one final test, the Shapiro-Wilk Normality test [57]. The Shapiro-Wilk test resulted in a *p-value* of 0.3449 and a *W* value of 0.9668. These results do not suggest rejection of the null hypothesis [21]

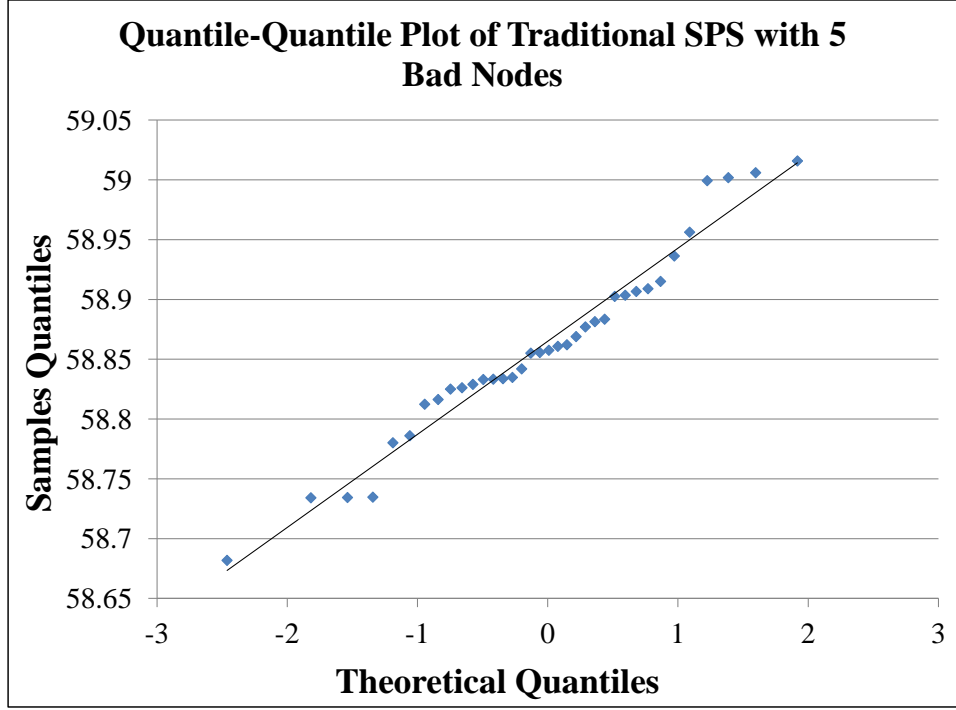


Figure 4.2: Quantile-Quantile Plot Generated from Simulation Data for Traditional SPS with 5 Bad Nodes

The selected 95% confidence interval corresponds to a statistical α value of 5%. Hence, a p -value less than α would cause us to reject the null hypothesis.

The W value is the ratio of the square of an approximate linear combination of sample ordered statistics by the symmetric estimate of variance. A large value close to 1 supports the null hypothesis.

As found in related research, both the p -value and w -value are sufficiently large, which indicates that the sample data was drawn from a Normally distributed population [21] [54]. The sample size of 36 observations, along with the results of the Histogram, Q-Q plot and Shapiro-Wilk test, further lend empirical in support of the *null hypothesis* that the sample came from Normally distributed data.

Since statistical results help confirm that the simulation environment generates Normal results, research can continue to Stage 2 to determine the optimal λ and Trust

Threshold values for the proposed modified exponentially weighted moving averages algorithm.

4.3 Stage 2 (Modified EWMA)

Stage 2 experiments were conducted in accordance with Methodology Section 3.6.2, in which the goal was to generate operating characteristic curve in which an optimal λ and Trust Threshold could be determined. The experiments were to completed utilizing modified EWMA Equations, 3.1 and 3.2, and success will be measured using properties of the receiver operating characteristic curve in Section 3.2.1. Figure 4.3 represents the result of this experiment.

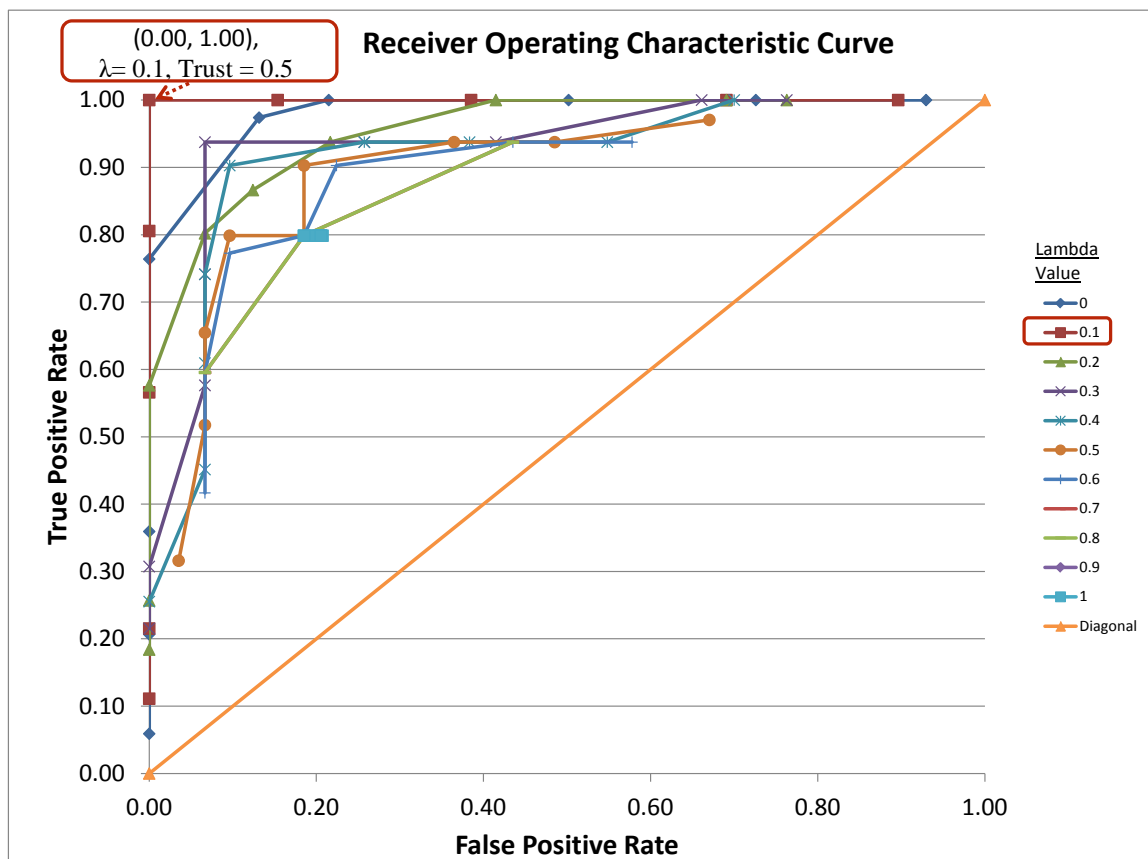


Figure 4.3: Receiver Operating Characteristic Curve Data with Optimal Strategy Identified; Point (0.00, 1.00) with $\lambda = 0.1$ and Trust Threshold = 0.5

A closer look at the data behind this point reveals the λ and Trust Threshold that produced point (0.00, 1.00) are 0.10 and 0.5 respectively, as observed in Figure 4.4. This point produces a special point on a receiver operating characteristic curve. The point (0.00, 1.00) in the top left corner denotes perfect classification: 100% true positive rate and 0% false positive rate [29]. Remaining simulations are conducted utilizing this optimal strategy.

A complete table of results is located in Appendix A, where all combinations of λ values ranging from 0.0 to 1.0 are calculated against Trust Thresholds ranging from 0.0 to 0.9, each in 0.1 increments.

		Trust Threshold									
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
Lambda Values	0.00	True Positives (TP)	576	576	576	576	561	440	207	119	34
		False Positives (FP)	502	392	271	116	71	0	0	0	0
		True Negatives (TN)	38	148	269	424	469	540	540	540	540
		False Negatives (FN)	0	0	0	0	15	136	369	457	542
		True Positive %	1.00	1.00	1.00	1.00	0.97	0.76	0.36	0.21	0.06
		False Positive %	0.93	0.73	0.50	0.21	0.13	0.00	0.00	0.00	0.00
	0.10	True Positives (TP)	576	576	576	576	576	464	326	124	64
		False Positives (FP)	484	373	208	83	0	0	0	0	0
		True Negatives (TN)	56	167	332	457	540	540	540	540	540
		False Negatives (FN)	0	0	0	0	0	112	250	452	512
		True Positive %	1.00	1.00	1.00	1.00	1.00	0.81	0.57	0.22	0.11
		False Positive %	0.90	0.69	0.39	0.15	0.00	0.00	0.00	0.00	0.00
	0.20	True Positives (TP)	576	576	576	540	499	462	332	148	106
		False Positives (FP)	412	373	224	117	67	36	0	0	0
		True Negatives (TN)	128	167	316	423	473	504	540	540	540
		False Negatives (FN)	0	0	0	36	77	114	244	428	470
		True Positive %	1.00	1.00	1.00	0.94	0.87	0.80	0.58	0.26	0.18
		False Positive %	0.76	0.69	0.41	0.22	0.12	0.07	0.00	0.00	0.00

Figure 4.4: Receiver Operating Characteristic Data With Optimal Strategy for λ and Trust Threshold Value Identified

In summary, although perfect classification is realized in this research, the ability of the system to maintain the minimum acceptable frequency of 58.8 Hz throughout each of the 5, 10 and 15 bad node test cases must still be validated in the next Stage.

4.4 Stage 3 (5, 10 and 15 Bad Node Frequencies)

Stage 3 experiments were conducted in accordance with Methodology Section 3.6.3, in which the goal was to determine the ability of the modified exponentially weighted moving averages algorithm, as compared to the special protection system without any trust management implemented. The modified EWMA is tested using the optimal λ and Trust Threshold values as determined in Section 4.3:

The goal of Stage 3 experiments is to determine the ability of the modified exponentially weighted moving averages algorithm, as compared to the special protection system without any trust management implemented, and related research that examines only the final trust value as a basis for trustworthiness (equal to $\lambda = 1$).

The experiments were completed utilizing modified EWMA Equations, 3.1 and 3.2, and the resulting optimal strategy that was identified in Figure 4.3. Figure 4.5 contains the results of the 15 bad node test case implemented with the optimal strategy.

At the 95% confidence interval, the modified EWMA trust module is able to maintain the frequency above the minimum acceptable frequency of 58.8 Hz during each individual simulation with the optimal strategy implemented. Conversely, the other two test cases the 15 bad node test case are not able to make the same claim. Figure 4.6 represents traditional reputation based trust approaches, which examine only the most recent trust observation to determine trustworthiness. This equates to a λ value of 1, as labeled in the graph. Upon further examination, figure 4.6 reveals that this approach to trust does achieve the required minimum acceptable frequency of 58.8 Hz some of the time with 95% confidence. However, the same confidence interval also reveals that a majority of the time, this approach also fails to meet the frequency requirement.

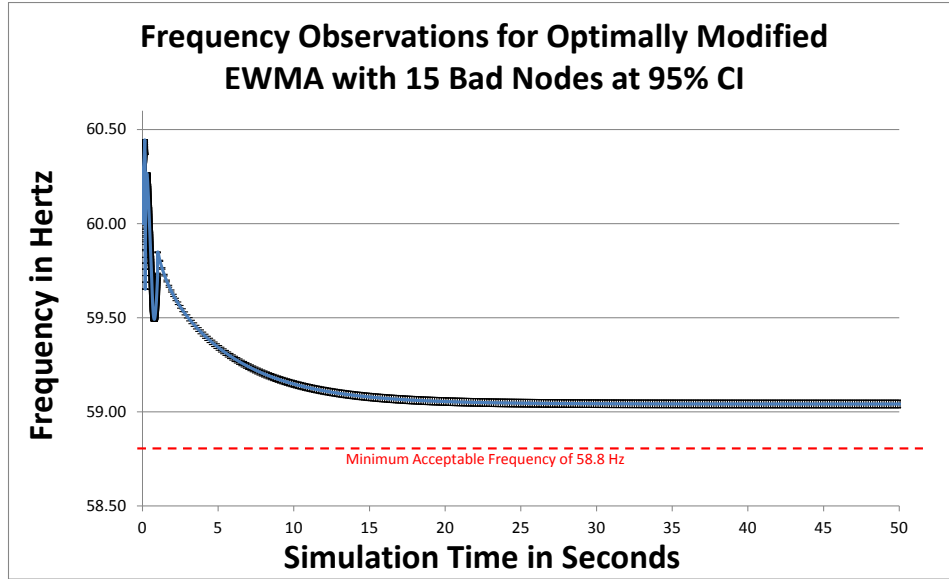


Figure 4.5: Frequency Observations for Optimally Modified EWMA with 15 Bad Nodes at 95% Confidence Interval

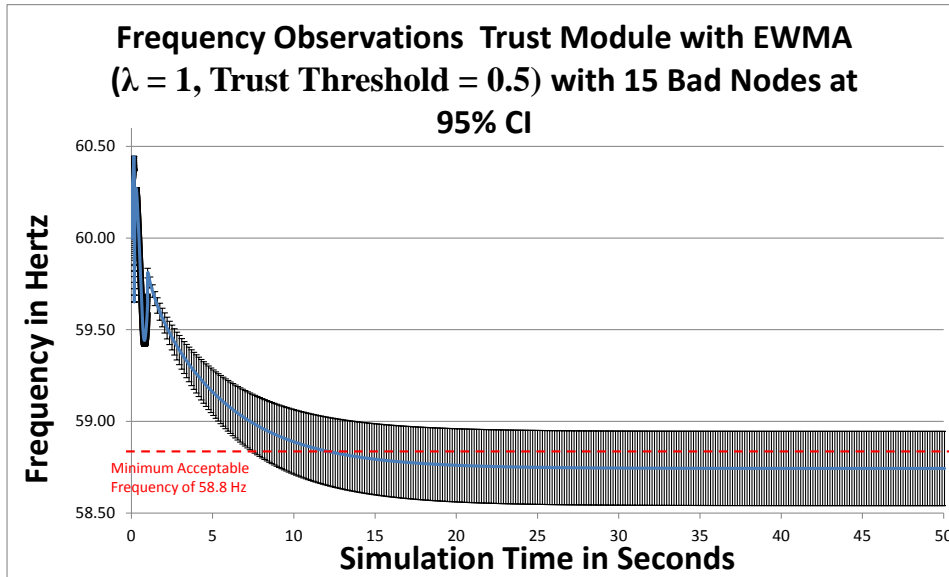


Figure 4.6: Frequency Observations for Traditional EWMA Trust Implementation ($\lambda = 1$, Trust Threshold = 0.5) with 15 Bad Nodes at 95% Confidence Interval

Finally, Figure 4.7 identifies the frequency observations associated with an implementation that does not utilize trust. As identified in the graph, with 95% confidence, each of the test cases can be expected to fail when there 15 bad nodes and no trust module is implemented.

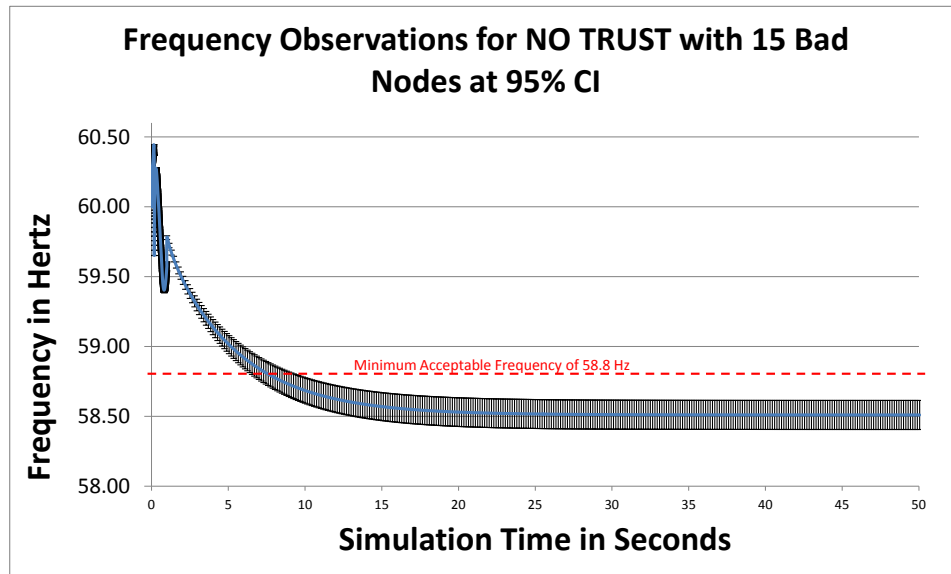


Figure 4.7: Frequency Observations for No Trust with 15 Bad Nodes at 95% Confidence Interval

Similar tests were conducted for both the 5 and 10 bad node test cases. The final frequency at the 95% confidence interval of each of these test cases are averaged and presented in Figure 4.8. As expected, each of the test cases with the modified EWMA trust module implemented maintains the frequency well above the minimum acceptable frequency. Conversely, with no trust module implemented, only the 5 bad node test case is able to meet the 58.8 Hz threshold.

Additionally, the Two-Factor Without Replication ANOVA Test analysis confirms the visual observations of this stage. Specifically, the test indicates a statistically significant difference in between trust module usage for $\lambda = 1$ and $\lambda = 0.1$ and/or non-trust usage for

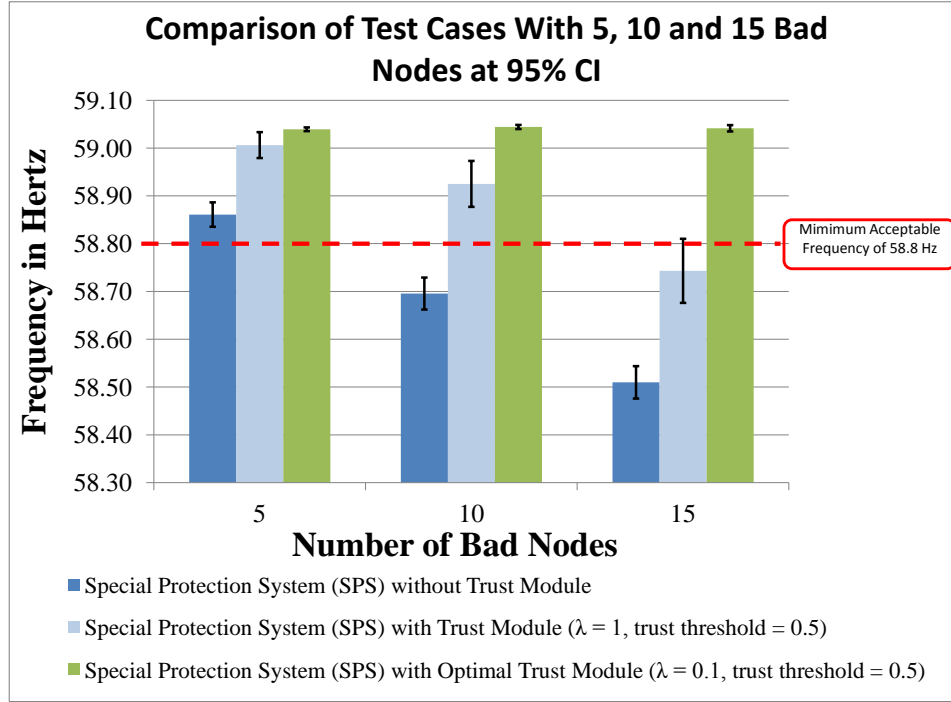


Figure 4.8: Comparison of Final Frequency Values for 5, 10 and 15 Bad Nodes at 95% Confidence Interval

each of the 5, 10 and 15 bad node test cases, as $f\text{-values} > f\text{-critical values}$ and $p\text{-values} < 0.05$ in each case. The cases that visually appear to have possible confidence interval overlap within Figure 4.8 are found within the 5 bad node cases that implement trust modules ($\lambda = 1$ and $\lambda = 0.1$ respectively). The ANOVA results for this analysis are $f\text{-value} = 5.925$, $f\text{-critical value} = 4.121$ and $p\text{-value} = 0.020$. At the 95% confidence interval, these ANOVA results are indicative of statistical difference.

One final ANOVA analysis of interest examined whether the optimal trust module strategies were statistically different between the 5, 10 and 15 bad node test cases, as the graph in Figure 4.8 would lead one to believe there is no difference. ANOVA analysis confirms this visual observation. The ANOVA results for this analysis are $f\text{-value} = 0.781$, $f\text{-critical value} = 3.128$ and $p\text{-value} = 0.462$. At the 95% confidence interval, these ANOVA results indicate a lack of statistical difference.

4.5 Analysis and Results Summary

In summary, the results of this research indicate that a modified exponentially weighted moving averages algorithm can successfully be applied to the trust module of a special protection system. In fact, between the 5, 10 and 15 bad node test cases, there was no statistical difference between the optimal trust strategy results. This was expected as the ROC curve in Figure 4.3 identified the strategy that ensured no false positives and no false negatives. Actual testing followed suit exactly. The simulation results fully support the use of the modified EWMA algorithm presented in this thesis for future smart grid special protections systems that implement reputation based trust.

5 Conclusion and Future Work

5.1 Chapter Overview

This chapter provides a conclusion of the work presented in this thesis. Just as important, this chapter also provides recommendation for future research. The results of the research in this thesis were so promising, that there are numerous follow-on experiments that can and should be done with regard to the modified exponentially weighted moving averages algorithm created herein.

5.2 Conclusions of Research

The main contribution of this thesis is the development and application of the modified Exponentially Weighted Moving Algorithm EWMA algorithm, and its ability to flawlessly function in the face varying numbers of bad (malicious or malfunctioning) special protection system nodes. This algorithm and its application contained herein should be implemented across current and future smart grid special protection system implementations.

Simulation results support the use of the proposed modified EWMA reputation based trust module in special protection systems within a smart grid environment. This modification resulted in the ability to maintain the associated frequency above the minimum acceptable frequency of 58.8 Hz in each of the 5, 10 and 15 *bad* node test cases at the 95% confidence interval. With regard to the modified EWMA algorithm itself, research concluded that the optimal λ and Trust Threshold create a trust module that is able to determine *good* nodes with a 100% true positive accuracy, and 0% false positive rate, resulting in a perfect classification scenario.

It is not the assertion of this researcher that the application of the modified EWMA algorithm to specific SPS architecture will create a perfect classification strategy, but

rather, that an optimal strategy will be revealed when the techniques described in this thesis are applied. This optimal strategy, by definition, promises to maximize true positives and minimize false positives.

Additionally, it is important to note that utilizing simulation frequencies and the method in which the "true" frequencies are determined was not a primal factor in this research. Rather, the frequencies, and how they were measured, are only representative of the type of data that can be imputed into the decision cycle of a reputation based trust special protection system. Future smart grid technologies and emerging SCADA intrusion detection technologies promise to increase the quantity of data available to make smarter trust calculations.

5.3 Recommendations for Future Research

Although not always attainable, a perfect classification for each specific application, as was realized in this research, is the goal for implementation at each field site. To this end, there are numerous recommendations for future research:

- As briefly touched upon in Section 5.2, there are numerous inputs that the smart grid promises to make available as possible inputs to the trust calculation of the future (e.g., demand response participation, amount of load drawn by the node and the nodes ability/willingness to contribute power in the event of an emergency).

These, as well as bolt-on technologies that have been created could serve as additional inputs to future iterations of this research. An example of one such input is a SCADA intrusion detection systems that monitors the SCADA network or system activities for malicious activities or policy violations. It is theorized that such inputs should weigh quite heavily as they could be alerting on an actual malicious event.

In addition to λ , these additional weighing factors can be included, and an optimal strategy can be subsequently determine based upon all inputs and their associated weights.

- The current simulation environment utilizes a window size of 16 to calculate the modified EWMA trust values. Due to the fact that data is continuously being generated and monitored within the smart grid, the actual window could be quite large. This large window promises to give an even better estimate of the actual trust value, as there is certainly more historical data on which to base the final trust value. To this end, additional simulations were conducted to see exactly where the current Optimal Trust value fails as depicted in Figure 5.1.

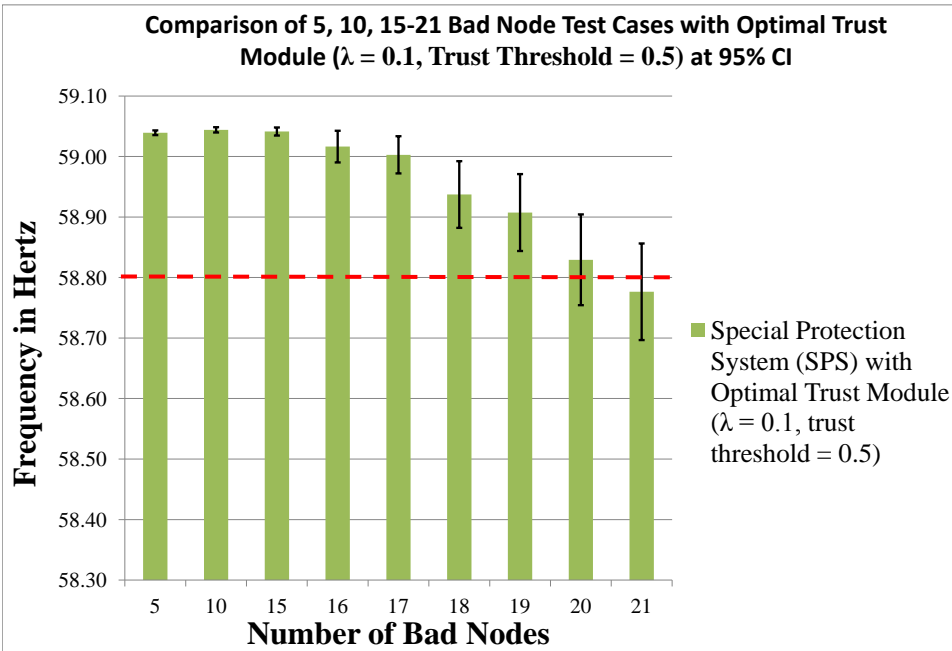


Figure 5.1: Comparison of 5, 10, 15-21 Bad Node Test Cases with Optimal Trust Module ($\lambda = 0.1$, Trust Threshold - 0.5) at 95% Confidence Interval

With the current window size, the most bad nodes that the system can manage, while ensuring that the minimum acceptable frequency is maintained with 95% confidence

is 19. Perhaps a larger window size will allow for a more bad-node-tolerant system. In exploring this facet of the algorithm, one must weight the cost of additional storage capacity for larger histories vs the desire to generate a more accurate trust.

- It is important to determine the robustness and appropriateness of the modified EWMA enhanced trust module across a variety of simulation environments. Numerous electrical power simulations exist, and the applicability of this algorithm to multiple testing environments will only lend credibly to its value and need for immediate real-world implementation.

In summary, implementation of a modified EWMA within a reputation based special protection system does account for each scenario that an electrical power engineer may face in the field. Instead, this research demonstrates that it provides a robust algorithm to incorporate within and test these challenges and/or opportunities upon.

Appendix: Receiver Operating Characteristic Data

Table A.1: Receiver Operating Characteristic Data Used to Determine Appropriate λ and Trust Threshold Values (1 of 2)

		Trust Threshold								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.00	True Positives (TP)	576	576	576	576	561	440	207	119	34
	False Positives (FP)	502	392	271	116	71	0	0	0	0
	True Negatives (TN)	38	148	269	424	469	540	540	540	540
	False Negatives (FN)	0	0	0	0	15	136	369	457	542
	True Positive %	1.00	1.00	1.00	1.00	0.97	0.76	0.36	0.21	0.06
	False Positive %	0.93	0.73	0.50	0.21	0.13	0.00	0.00	0.00	0.00
0.10	True Positives (TP)	576	576	576	576	576	464	326	124	64
	False Positives (FP)	484	373	208	83	0	0	0	0	0
	True Negatives (TN)	56	167	332	457	540	540	540	540	540
	False Negatives (FN)	0	0	0	0	0	112	250	452	512
	True Positive %	1.00	1.00	1.00	1.00	1.00	0.81	0.57	0.22	0.11
	False Positive %	0.90	0.69	0.39	0.15	0.00	0.00	0.00	0.00	0.00
0.20	True Positives (TP)	576	576	576	540	499	462	332	148	106
	False Positives (FP)	412	373	224	117	67	36	0	0	0
	True Negatives (TN)	128	167	316	423	473	504	540	540	540
	False Negatives (FN)	0	0	0	36	77	114	244	428	470
	True Positive %	1.00	1.00	1.00	0.94	0.87	0.80	0.58	0.26	0.18
	False Positive %	0.76	0.69	0.41	0.22	0.12	0.07	0.00	0.00	0.00
0.30	True Positives (TP)	576	576	540	540	540	427	332	177	124
	False Positives (FP)	412	357	224	139	36	36	36	0	0
	True Negatives (TN)	128	183	316	401	504	504	504	540	540
	False Negatives (FN)	0	0	36	36	36	149	244	399	452
	True Positive %	1.00	1.00	0.94	0.94	0.94	0.74	0.58	0.31	0.22
	False Positive %	0.76	0.66	0.41	0.26	0.07	0.07	0.07	0.00	0.00
0.40	True Positives (TP)	576	540	540	540	520	427	351	260	147
	False Positives (FP)	378	296	207	139	52	36	36	36	0
	True Negatives (TN)	162	244	333	401	488	504	504	504	540
	False Negatives (FN)	0	36	36	36	56	149	225	316	429
	True Positive %	1.00	0.94	0.94	0.94	0.90	0.74	0.61	0.45	0.26
	False Positive %	0.70	0.55	0.38	0.26	0.10	0.07	0.07	0.07	0.00
0.50	True Positives (TP)	559	540	540	520	460	460	377	298	182
	False Positives (FP)	362	262	197	100	100	52	36	36	19
	True Negatives (TN)	178	278	343	440	440	488	504	504	521
	False Negatives (FN)	17	36	36	56	116	116	199	278	394
	True Positive %	0.97	0.94	0.94	0.90	0.80	0.80	0.65	0.52	0.32
	False Positive %	0.67	0.49	0.36	0.19	0.19	0.10	0.07	0.07	0.04

Table A.2: Receiver Operating Characteristic Data Used to Determine Appropriate λ and Trust Threshold Values (2 of 2)

		Trust Threshold									
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
Lambda Values	0.60	True Positives (TP)	540	540	520	460	460	460	445	343	240
		False Positives (FP)	312	235	121	100	100	100	52	36	36
		True Negatives (TN)	228	305	419	440	440	440	488	504	504
		False Negatives (FN)	36	36	56	116	116	116	131	233	336
		True Positive %	0.94	0.94	0.90	0.80	0.80	0.80	0.77	0.60	0.42
		False Positive %	0.58	0.44	0.22	0.19	0.19	0.19	0.10	0.07	0.07
	0.70	True Positives (TP)	540	540	460	460	460	460	460	343	343
		False Positives (FP)	235	235	100	100	100	100	100	36	36
		True Negatives (TN)	305	305	440	440	440	440	440	504	504
		False Negatives (FN)	36	36	116	116	116	116	116	233	233
		True Positive %	0.94	0.94	0.80	0.80	0.80	0.80	0.80	0.60	0.60
		False Positive %	0.44	0.44	0.19	0.19	0.19	0.19	0.19	0.07	0.07
	0.80	True Positives (TP)	540	460	460	460	460	460	460	460	343
		False Positives (FP)	235	100	100	100	100	100	100	100	36
		True Negatives (TN)	305	440	440	440	440	440	440	440	504
		False Negatives (FN)	36	116	116	116	116	116	116	116	233
		True Positive %	0.94	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.60
		False Positive %	0.44	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.07
	0.90	True Positives (TP)	460	460	460	460	460	460	460	460	460
		False Positives (FP)	100	100	100	100	100	100	100	100	100
True Negatives (TN)		440	440	440	440	440	440	440	440	440	
False Negatives (FN)		116	116	116	116	116	116	116	116	116	
True Positive %		0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	
False Positive %		0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	
1.00	True Positives (TP)	460	460	460	460	460	460	460	460	460	
	False Positives (FP)	111	111	112	107	106	111	111	100	100	
	True Negatives (TN)	429	429	428	433	434	429	429	440	440	
	False Negatives (FN)	116	116	116	116	116	116	116	116	116	
	True Positive %	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	
	False Positive %	0.21	0.21	0.21	0.20	0.20	0.21	0.21	0.19	0.19	

Bibliography

- [1] “Guide for Abnormal Frequency Protection for Power Generating Plants”. *IEEE Std C37.106-2003*, 1–34, 2004.
- [2] Abdul-Rahman, A. and S. Hailes. “Supporting trust in virtual communities”. *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*. IEEE, 2000.
- [3] Anderson, P.M. *Power system protection*. McGraw-Hill New York, 1999.
- [4] Anderson, PM and BK LeReverend. “Industry experience with special protection schemes”. *IEEE Transactions on Power Systems*, 11(3):1166–1179, 1996.
- [5] Azzedin, F. and M. Maheswaran. “Evolving and managing trust in grid computing systems”. *Canadian Conference on Electrical and Computer Engineering*, volume 3, 1424–1429. IEEE, 2002.
- [6] Barker, P.P. and RW De Mello. “Determining the impact of distributed generation on power systems. I. Radial distribution systems”. *Power Engineering Society Summer Meeting*, volume 3, 1645–1656. IEEE, 2000.
- [7] Barsom, J.M. and S.T. Rolfe. *Fracture and fatigue control in structures: Applications of fracture mechanics*, volume 41. ASTM International, 1999.
- [8] Bengiamin, NN and WC Chan. “Variable structure control of electric power generation”. *Power Apparatus and Systems, IEEE Transactions on*, (2):376–380, 1982.
- [9] Blaze, M., J. Feigenbaum, and J. Lacy. “Decentralized trust management”. *IEEE Symposium on Security and Privacy*, 164–173. 1996.
- [10] Borowski, J.F. *Reputation-Based Trust for a Cooperative, Agent-Based Backup Protection Scheme for Power Networks*. Technical report, DTIC Document, 2010.
- [11] Bowen III, C.L., T.K. Buennemeyer, and R.W. Thomas. “Next generation SCADA security: best practices and client puzzles”. *Proceedings from the Sixth Annual IEEE SMC Information Assurance Workshop*, 426–427. 2005.
- [12] Box, G.E.P., G.M. Jenkins, and J.F. MacGregor. “Some recent advances in forecasting and control”. *Applied Statistics*, 158–179, 1974.
- [13] Clinton, W.J. “Executive order 13010-critical infrastructure protection”. *Federal Register*, 61(138):37347–37350, 1996.
- [14] Clinton, W.J. “Presidential Decision Directive 63”. *Government Printing Office, Washington, DC*, 1998.

- [15] Coates, G.M., K.M. Hopkinson, S.R. Graham, and S.H. Kurkowski. "Collaborative, trust-based security mechanisms for a regional utility intranet". *IEEE Transactions on Power Systems*, 23(3):831–844, 2008.
- [16] Coates, G.M., K.M. Hopkinson, S.R. Graham, and S.H. Kurkowski. "A trust system architecture for SCADA network security". *IEEE Transactions on Power Delivery*, 25(1):158–169, 2010.
- [17] Commission, Federal Energy Regulatory. "Smart grid policy". *Docket No. PL09-4-000*.
- [18] Daneels, A. and W. Salter. "What is SCADA". *International Conference on Accelerator and Large Experimental Physics Control Systems*, 339–343. 1999.
- [19] Egan, JP. "Signal Detection Theory and ROC Analysis, Series in Cognition and Perception". 1975.
- [20] Fadul, J., K. Hopkinson, C. Sheffield, J. Moore, and T. Andel. "Trust Management and Security in the Future Communication-Based Smart Electric Power Grid". *2011 44th Hawaii International Conference on System Sciences (HICSS)*, 1–10. IEEE, 2011.
- [21] Fadul, J.E. *Using Reputation Based Trust to Overcome Malfunctions and Malicious Failures in Electric Power Protection Systems*. Ph.D. thesis, 2011.
- [22] Fawcett, T. "An introduction to ROC analysis". *Pattern recognition letters*, 27(8):861–874, 2006.
- [23] Fernandez, J.D. and A.E. Fernandez. "SCADA systems: vulnerabilities and remediation". *Journal of Computing Sciences in Colleges*, 20(4):160–168, 2005.
- [24] Giani, A., G. Karsai, T. Roosta, A. Shah, B. Sinopoli, and J. Wiley. "A testbed for secure and robust SCADA systems". *ACM SIGBED Review*, 5(2):4, 2008.
- [25] Gorman, S. "Electricity grid in US penetrated by spies". *Wall Street Journal*, 8, 2009.
- [26] Graham, R. and D. Maynor. "SCADA security and terrorism: Were not crying wolf". *Black Hat Federal*, Washington, D.C., 2006.
- [27] Grandison, T. and M. Sloman. "A survey of trust in internet applications". *IEEE Communications Surveys & Tutorials*, 3(4):2–16, 2000.
- [28] Grimes, M. "SCADA exposed". *Proc. ToorCon*, 7, 2005.
- [29] Hamel, L. "Model assessment with ROC curves". *The Encyclopedia of Data Warehousing and Mining*.

- [30] Holt, C.C. “Forecasting seasonals and trends by exponentially weighted moving averages”. *International Journal of Forecasting*, 20(1):5–10, 2004.
- [31] Hopkinson, K., G. Roberts, X. Wang, and J. Thorp. “Quality-of-service considerations in utility communication networks”. *IEEE Transactions on Power Delivery*, 24(3):1465–1474, 2009.
- [32] Hopkinson, K., X. Wang, R. Giovanini, J. Thorp, K. Birman, and D. Coury. “EPOCHS: a platform for agent-based electric power and communication simulation built from commercial off-the-shelf components”. *Power Systems, IEEE Transactions on*, 21(2):548–558, 2006.
- [33] Hunter, J.S. “The exponentially weighted moving average.” *J. QUALITY TECHNOL.*, 18(4):203–210, 1986.
- [34] Ijure, V.M., S.A. Laughter, and R.D. Williams. “Security issues in SCADA networks”. *Computers & Security*, 25(7):498–506, 2006.
- [35] Issariyakul, T. and E. Hossain. *Introduction to network simulator NS2*. Springer Verlag, 2011.
- [36] Khurana, H., M. Hadley, N. Lu, and D.A. Frincke. “Smart-grid security issues”. *IEEE Security & Privacy*, 8(1):81–85, 2010.
- [37] Kintner-Meyer, M., K. Schneider, and R. Pratt. “Impacts assessment of plug-in hybrid vehicles on electric utilities and regional US power grids, Part 1: Technical analysis”. *Pacific Northwest National Laboratory*, 2007.
- [38] Klinger, M., WA Mittelstadt, and CW Taylor. “Transient Stability Controls Used by Bonneville Power Administration to Mitigate Delays of Planned Facilities”. *CIGRE, Paris*, 1982.
- [39] Kundur, P., N.J. Balu, and M.G. Lauby. *Power system stability and control*, volume 4. McGraw-hill New York, 1994.
- [40] LeBlanc, D.C. *Statistics: concepts and applications for science*. Jones & Bartlett Pub, 2004.
- [41] Lewis, J.E. “The Economic Espionage Act and the Threat of Chinese Espionage in the United States”. *J. Intell. Prop.*, 8:189, 2008.
- [42] Liscouski, B. and W. Elliot. “Final report on the august 14, 2003 blackout in the united states and canada: Causes and recommendations”. *A report to US Department of Energy*, 40, 2004.
- [43] Lucas, J.M. and M.S. Saccucci. “Exponentially weighted moving average control schemes: properties and enhancements”. *Technometrics*, 1–12, 1990.

- [44] Marsh, S. “Trust and reliance in multi-agent systems: A preliminary report”. *Proceedings of Jth European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, 94–112. Citeseer, 1992.
- [45] Matus, P.A. *Strategic Impact of Cyber Warfare Rules for the United States*. Technical report, DTIC Document, 2010.
- [46] Misztal, B.A. “Trust in modern societies. The search for the bases of social order”. *Cambridge et al*, 1996.
- [47] Moslehi, K. and R. Kumar. “Smart grid-a reliability perspective”. *Innovative Smart Grid Technologies (ISGT)*, 1–8. 2010.
- [48] Muth, J.F. “Optimal properties of exponentially weighted forecasts”. *Journal of the american statistical association*, 299–306, 1960.
- [49] NSTAC, Information Assurance Task Force. *Electric Power Information Assurance Risk Assessment*. Technical report, NSTAC Document, March 1997.
- [50] NSTAC, The President’s National Security Telecommunicaitons Advisory Committee. *The NSTAC’s Response to the National Plan*. Technical report, NSTAC Document, April 2001.
- [51] Padiyar, K.R. *Power system dynamics-stability and control*. 2. B.S. Publications, 2002.
- [52] Powner, D.A. *Electricity Grid Modernization: Progress Being Made on Cybersecurity Guidelines, But Key Challenges Remain to be Addressed*. DIANE Publishing, 2011.
- [53] Roberts, SW. “Control chart tests based on geometric moving averages”. *Technometrics*, 239–250, 1959.
- [54] Ross, K.J. *Application of Game Thoery to Improve the Defense of the Smart Grid*. Technical report, DTIC Document, 2012.
- [55] Sankar, P.B. and C.S. Babu. “Transient stability enhancement of power system using STATCOM”. *Inter. Jour. of Elec. and Pow. Engi*, (2):271–276, 2008.
- [56] Schmid, W. and A. Schone. “Some properties of the EWMA control chart in the presence of autocorrelation”. *The Annals of Statistics*, 1277–1283, 1997.
- [57] Shapiro, S.S., M.B. Wilk, and H.J. Chen. “A comparative study of various tests for normality”. *Journal of the American Statistical Association*, 1343–1372, 1968.
- [58] Shaw, W. *Cybersecurity for SCADA systems*. PennWell Corporation, 5 edition, 2006. ISBN 1-59370-068-7.

- [59] Sheffrin, A., H. Yoshimura, D. LaPlante, and B. Neenan. “Harnessing the power of demand”. *The Electricity Journal*, 21(2):39–50, 2008.
- [60] Siemens Energy. “PSS/E Product Suite”, 2012. URL <http://www.energy.siemens.com/us/en/services/power-transmission-distribution/power-technologies-international/software-solutions/pss-e.htm>.
- [61] Sitts, G. “Radio pierces the great blackout”. *Broadcast Engineering*.
- [62] Spackman, K.A. “Signal detection theory: Valuable tools for evaluating inductive learning”. *Proceedings of the sixth international workshop on Machine learning*, 160–163. Morgan Kaufmann Publishers Inc., 1989.
- [63] Strobel, C.D. “American recovery and reinvestment act of 2009”. *Journal of Corporate Accounting & Finance*, 20(5):83–85, 2009.
- [64] Vittal, V. “Transient stability test systems for direct stability methods”. *IEEE Transactions on Power Systems (Institute of Electrical and Electronics Engineers);(United States)*, 7(1), 1992.
- [65] Wilshusen, G.C. and D.A. Powner. *CYBERSECURITY: Continued Efforts Are Needed to Protect Information Systems from Evolving Threats*. Technical report, DTIC Document, 2009.
- [66] Winter, W. and B. LeReverend. “Operational performance of bulk electricity system control aids”. *Electra*, 123:7–101, 1989.
- [67] Xie, Z., G. Manimaran, V. Vittal, AG Phadke, and V. Centeno. “An information architecture for future power systems and its reliability analysis”. *IEEE Transactions on Power Systems*, 17(3):857–863, 2002.
- [68] Xiong, L. and L. Liu. “A reputation-based trust model for peer-to-peer e-commerce communities”. *IEEE International Conference on E-Commerce*, 275–284. 2003.
- [69] Zhu, Z., S. Zhao, J.D. McCalley, V. Vittal, and AA Irizarry-Rivera. “Risk-based security assessment influenced by generator rejection”. *Proceedings of the Sixth Probabilistic Methods Applied to Power Systems (PMAPS) International Conference*. 1997.

Vita

Captain Andrew Kasperek is a student at the Air Force Institute of Technology pursuing a Masters Degree in Cyber Operations. He graduated from Clemson University with the academic degree of Bachelor of Science in Computer Science in 2006. He also obtained a Masters of Information Systems degree from the University of Phoenix in 2007. Captain Kasperek is a member of the Upsilon Pi Epsilon honor society.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 074-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) 14-06-2012		2. REPORT TYPE Master's Thesis		3. DATES COVERED (From – To) August 2010 – June 2012	
4. TITLE AND SUBTITLE Enhancing Trust in the Smart Grid by Applying a Modified Exponentially Weighted Averages Algorithm				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Kasperek, Andrew, T., Capt				5d. PROJECT NUMBER 12G292P	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865				8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/GCO/ENG/12-18	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Force Office of Scientific Research, Mathematics, Information and Life Sciences Directorate Attn : Dr. Robert J. Bonneau 875 N Randolph St, Ste 325, Rm 3112, Arlington, VA 22203 (703) 696-9545 (DSN: 426-9545) Email: robert.bonneau@afosr.af.mil				10. SPONSOR/MONITOR'S ACRONYM(S) AFOSR/NL	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED					
13. SUPPLEMENTARY NOTES This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States					
14. ABSTRACT The main contribution of this thesis is the development and application of a modified Exponentially Weighted Moving Algorithm (EWMA) algorithm, and its ability to robustly function in the face varying numbers of bad (malicious or malfunctioning) Special Protection System (SPS) nodes. Simulation results support the use of the proposed modified EWMA reputation based trust module in SPSs within a smart grid environment. This modification results in the ability to easily maintain the system above the minimum acceptable frequency of 58.8 Hz at the 95% confidence interval, when challenged with test cases containing 5, 10 and 15 bad node test cases out of 31 total load nodes. These promising results are realized by incorporating the optimal modified EWMA strategy, as identified by Receiver Operating Characteristic (ROC) techniques, where an optimal strategy is revealed. The optimal strategy maximizes true positives while minimizing false positives. Implementation of a modified EWMA within a reputation based special protection system does not account for each scenario that an electrical power engineer may face in the field. Instead, this research demonstrates that such an algorithm provides a robust environment to test within, in the hope of successfully meeting challenges and/or opportunities of the future.					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:		17. LIMITATION OF ABSTRACT		18. NUMBER OF PAGES	
REPORT U	ABSTRACT U	c. THIS PAGE U	UU	86	19a. NAME OF RESPONSIBLE PERSON Kenneth M. Hopkinson, Civ, USAF (ENG)
					19b. TELEPHONE NUMBER (Include area code) (937) 255-6565, ext 4579 (Kenneth.hopkinson@afit.edu)